1965 CONGRESS

INTERNATIONAL FEDERATION FOR DOCUMENTATION (FID)

WASHINGTON, D. C.
U. S. A.

10-15 OCTOBER 1965

Abstracts

1965 Congress

International Federation for Documentation (FID)

10 - 15 OCTOBER 1965

WASHINGTON, D. C.

Abstracts

SECRETARIAT, 1965 FID CONGRESS

9650 WISCONSIN AVENUE, WASHINGTON, D C 20014, U.S.A.

CONTENTS

This booklet contains abstracts of both symposium papers and contributed papers, followed by an index of authors. Symposium abstracts are grouped sequentially under the five Area topics listed above, and appear in the order in which the papers will be presented at the Congress. Contributed abstracts are ordered alphabetically by the name of the first author. Numerals following author names in the index refer to the page number of the abstract.

AREA I. EDUCATION AND TRAINING FOR DOCUMENTALISTS

Symposium A: Present State of Education and Training

THE TRENDS OF DOCUMENTALIST TRAINING IN EAST EUROPEAN COUNTRIES. Zygmunt Majewski. Committee for Science and Technology, Warsaw, Krakowskie Przedmiescie 1, Poland.

I understand the East European countries to be the socialist countries that mainly occupy this part of Europe.

The economic system which is specific for these countries is projected in a considerable measure onto other aspects of national life. It is projected onto scientific informational activities also in spite of the fact that it utilizes in its work methods and measures applied in other countries of the world.

The national economy in socialist countries is nationalized and the industry, communication equipment, and other economic branches are the property of the entire community and are centrally directed by appropriate economic organizations. The general outline of all the functions, production etc. are planned as a whole for several years in advance and the plans are expressed in respective state acts. Certain wide differences emerge among particular countries as regards agriculture which is sometimes wholly nationalized or just partly /as in Poland, for instance/.

Scientific institutions also have their research plans. Such plans are coordinated by one state body irrespective of whom the given institution is subordinate to./ Academy of Sciences - Higher Schools of Learning - Economic Departments/.

The influence of such a scientific and economic system on the work of the scientific information service is so reflected that there are no secrets between work establishments and scientific institutions and information can flow freely. Having centralized managements of entire branches of the national economy, central occupational information services can easily be formed. The same concerns the branches of the basic sciences. Having plans for the development of learning and economy for the future years, information centers can prepare in good time the necessary scientific information and be in a certain measure the coauthors of the plans mentioned.

How does documentalist training look and to what extent does the state of affairs delineated above reflect on documentalist training?

Please note that the question here is the training of documentalists exclusively those who are university graduates and not those who are merely graduates of a documentalist technical school /graduates of secondary schools and a professional documentational school as, for instance in Poland or assistant personnel-laboratory workers, for instance/.

Training university graduates was conducted until now at regular attendance courses not exceeding 100 lecture hours and exercises or at extramural courses of longer duration. Outside of a "classical" curriculum for documentation and scientific information these curricula took into greater consideration subjects facilitating cooperation between scientific information services and scientific or

economic institutions. Such subjects cover the organization of science and economy, cooperation between economy and science, planning methods etc. In as much as computers are used with increasing frequency both in planning and scientific information, the principles of their operation are one of the important subjects. Analytical and synthetic reports are a means of information used with increasing frequency; writing up such reports is also a subject of lectures. The same concerns using direct information /not by way of the document/ such as organizing scientific conferences, radio broadcasts and television.

Almost all the East European countries have prepared postgraduate curricula generally for one year of about 1000 hours of lectures and exercises. Such studies are arranged by colleges and universities which will be able to confer doctor's degrees.

Of the socialist countries only the USSR confers doctor's degrees in scientific information.

Training of users of documentation and scientific information is conducted in all the socialist countries both in colleges and universities /15 to 36 hours/ and at specialistic courses, as for instance training courses for managers.

EDUCATION AND TRAINING OF DOCUMENTALIST. A. R. Mohajir. Pakistan National Science and Technical Documentation Center. Karachi, Pakistan.

Various problems involved in the education and training of persons required in different fields of documentation and information service are reviewed at length. Educational requirements, knowledge and training background in documentation necessitated by general, scientific, industrial documentation; documentation in arts, social sciences, literature and humanities; information management; economic and managerial information; translations and document reproduction have been exhaustively and systematically enumerated. A person with a basic degree would be more useful for documentation purposes than a specialist of a particular field. Library science qualifications or library training would be an additional advantage. Appropriate educational background with temperamental qualities would be more effective in the operation of the documentation services. To meet with the situation education in documentation at University level or by special institutions has been recommended. Training of documentalists by FID in its future program is also cross referenced. Training of the documentation personnel on Job basis, formal training, training in the home country or in foreign countries have been discussed in detail. Since training facilities in developing countries of Asia are very limited, foreign training is therefore highly desirable to improve the standards of documentation service. The role of Pakistan National Scientific & Technical Documentation Center (PANSDOC) with special reference to its training facilities offered to scientific and technical institutions and industries and providing courses on scientific and technical documentation to University students at post-graduate level have been briefly mentioned. It is suggested that a detailed study and analysis of the observations, experiences and results accrued from the existing systems of education and training in documentation would be conducive to improve and standardize the methods, systems, syllabi and practices regarding the education and training in documentation. Status of documentalists is evaluated from professional and social point of view; and it is pointed out that the documentalists have vital roles in the country's progress and hence their status should receive due recognition. In Pakistan, position and status of documentalists are at par with the scientific personnel.

THE PRESENT STATE OF EDUCATION AND TRAINING IN DOCUMENTATION, INFORMATION SCIENCE, AND SPECIAL LIBRARIANSHIP. Jesse H. Shera. School of Library Science, Western Reserve University, Cleveland, Ohio, U. S. A.

The origins of contemporary thought concerning the professional education of documentalists and other specialists in information science are probably to be found in speculation about the need for an adequate educational program for the professional preparation of special librarians. As early as 1911, only two years after the Special Libraries Association was established, A. G. S. Josephson of the John Crerar Library pointed out in a letter to the Special Libraries Association the need for specialization in library education. Seven years later, in 1918, J. G. Pearce argued in the pages of the British Library Association Record for training in documentation as a special branch of library education in general.

Thereafter, most of the argument concerning specialization in librarianship came from the special librarians, and little attention was paid to the field known today as information science. In 1926, Linda Morley offered at Columbia University the first course in special librarianship to be sponsored by any library school, and similar courses were slowly introduced into other library schools.

On this side of the Atlantic documentation was, during the 1930's, largely interpreted in terms of microphotographic and other photographic processes, and the concept was not extended to include non-conventional methods of information storage and retrieval until after the end of the Second World War. In 1950 Helen Focke of the School of Library Science at Western Reserve University was the first to offer a course in documentation, and five years later the program was greatly expanded with the establishment of the Center for Documentation and Communication Research as a division of the library school at Western Reserve.

The argument whether a special librarian should be trained primarily as a subject specialist or as a librarian was quickly transferred to the discussion of the professional education of the documentalists and information specialists. A study by Cohen and Craven (financed by the National Science Foundation) found from their sampling of opinions of practicing information specialists that such conventional subjects as information sources, cataloging and classification, and subject bibliography were given a high priority, but that computer technology, information science, and mechanized information techniques ranked low on the list. A number of conferences have failed really to clarify the problem of professional education, much less develop any authoritative conclusions or evolve a consensus.

At the present time the professional education of documentalists and information specialists in the United States falls into two major groups, those attached to conventional library schools, as at Western Reserve, the University of Minnesota, and the Drexel Institute of Technology, and those which are not based upon librarianship at all as at the Georgia Institute of Technology, Lehigh University, and, most recently, at the University of North Carolina. In a few instances special degrees in information science have been established, but as yet none has been recognized by any authoritative accrediting agency.

There are many unresolved problems which must be solved before a definitive program, or programs, can be established. The field itself must achieve a higher degree of standardization, and its terminology must be stabilized. Effective studies of the intellectual needs and skills of successful practitioners of

documentation are desperately needed. All of the programs are, at present, too young to have produced graduates whose competence has been tested over a sufficiently long period of time. Programs have been predicated on assumptions of what is and should be rather than on any precise knowledge of what the information specialist does and needs to know in order to do it. The role of automation and computer technology in this newly emerging field of documentation is only vaguely understood, if indeed it is understood at all. Much cautious and well-founded exploratory work and experimentation must be done if the training of a substantial, perhaps even disastrous, number of misfits is to be avoided.

TRAINING OF DOCUMENTALISTS IN SOUTH AMERICA. Roberto Couture de Troismonts. Fundación Interamericana de Bibliotecología Franklin, Avda. del Libertador 1146, Buenos Aires, Argentina.

One of the most remarkable consequences of progress, both technical and scientific is an immediate production of "documents" in all their forms - intellectual and physical - and the problem of preparation, classification, retrieval and locating of the information contained in them. This phenomenon, in a greater or lesser degree, directly or indirectly, happens in every country in the world and has aroused a demand for specialists in the use of documentary techniques. As is well known the cultural, economic, social, scientific and technical development of the Latin American countries included approximately between latitude north 30º and latitude south 50º is unequal. Undoubtedly these aspects should be studied in their relation to the continent as a whole but, in every case taking into account the individual conditions and necessities of each country.

To introduce the teaching of documentation it is essential to know beforehand the most urgent requirement that would contribute to the industrial development or the solution of the economic, social, educational or health problems (1). We know that courses in documentation are held in Argentina (Curso regional de orientación a la documentación científica y técnica-UNESCO-C.N.I.C. y T.), Brazil (I.B.B.D.) and Chile (CENID) and according to a recent report from the Medellin Seminar on Library Science studies, we know that out of 33 schools, 12 include this subject. The programs of study are not sufficient to give us a clear conception of what the classes comprise. As Gaston Julia used to say: "What is important are the ideas more than the subjects which are taught, the manner of working more than the deeds." The great majority of librarians in Latin America - as in other more developed countries - have not had the opportunity of acquiring a documentary conscience, a fact which would have no importance if they would limit themselves to the tipical librarian function of an "educational" nature (2); but due to the lack of human resources they must face the organizing of documentation centers, special and specialized services of information, etc. It is not worthwhile to repeat here the sterile discussions which occur between librarians and documentalists, let us say however, that there is a new profession which is no more than ten years old and which we shall call "A", which must tend to such a variety of requirements; as some of the ones which are mentioned here: preparation of the documentation and automatic processing of information in public administration (3); (in most countries the same or almost the same limitations and defficiences can be found which affect the adequate functioning of the administrative machinery as an instrument of public service and general progress) (4); statistics, census and sociological surveys, scientific and technical documentation of: firms, banks and insurance companies, market surveys, registration of persons, register of patents, legislative and fiscal, information for the armed forces, newspaper and publishing enterprises (as much for the information they gather as for the means of dissemination) (5); if we put aside the principle that documentation is part of library science, inasmuch as, besides having different methods and techniques, it is a mental attitude, it will be easier to determine the characteristics of the professional training. In Latin America undoubtedly there exist common problems, but it should be established which are the most urgent necessities of each country and specify what appointment will be given to the trained people (public administration, type of industry, research center).

We should insist that the documentalists as well as the librarians should undertake activities eminently practical and not speculative. As a result of a recent trip through various Latin American countries (Chile, Peru, Colombia, Venezuela,

Mexico and Uruguay) we gathered the impression that in those countries there exists interest for courses in documentation. At this time it does not seem opportune to found schools for long term studies; it would be sufficient to establish the following courses, choosing those most apt according to the country and the circumstances: A) Lectures or short elementary courses of information sponsored by scientific institutions, school for librarians, professional associations, documentation centers, universities...; B) Higher level courses of from 4 to 6 months for young researchers, engineers, technicians, post-graduates, government officials in technical posts, graduates from higher level schools for librarians. The course could be divided into 1) general documentation and 2) specialized documentation, which in turn could be subdivided in: a) social, economic and humanistic sciences and b) pure and applied science, technology. These courses would be sponsored by Councils or National Centers of research or documentation or higher institutes that play the role of receiver of information problems. It would be adviseable to be able to count for these classes with the participation of professors of international repute in order that the very latest advances in the field of documentation be included as well as experimental work and prospects for the future. C) Complementary courses and/or documentary subjects included in the program of study in schools for librarians.

The basic formation we have outlined could be used for distinct categories of personnel which the services require, some of which have been mentioned by Koblitz (6). The furthering of the courses in each country could be accomplished through the Secretary of FID/CLA/.

(1) América en cifras, 1963, v.3.t.2. Industria. Washington, Union Panamericana, 1964.
--- v.5. Situación cultural.
(2) Shera, Jesse H. y Egan, Margaret E. "Examen del estado actual de la biblioteconomía y de la documentación". En Revista Universidad, Universidad nacional del Litoral (Santa Fe), nº 61, 1964. p. 279-321.
(3) Sadosky, Manuel y Gianetto, Guillermina G. "La documentación científica y la automatización". En Seminario latinoamericano sobre documentación científica, Lima, 3-8 setiembre 1962 (Montevideo). Centro de cooperación científica de la Unesco para la América Latina, 1962.
(4) "La administración pública en América Latina". Resumen de la publicación OEA/ser. H/X4-CIES/348/español, 27 set. 63. En Revista de administración pública, (Buenos Aires), nº 12, mar. 1964. p. 80-118.
(5) Taube, Mortimer and others, Studies in coordinate indexing /s.l./ Documentation Incorporated, 1953, v.1.
(6) Koblitz, Josef. "Factores humanos", en Otto Frank y otros, Técnicas modernas de documentación e información. Buenos Aires, Eudeba, 1964. p. 152-170.

A SWEDISH COURSE IN SCIENTIFIC INFORMATION TRANSFER AND RETRIEVAL. Björn V. Tell. Royal Institute of Technology, Stockholm 70, Sweden.

During the Fall and Spring semesters 1964/65 a course called "Transmission of Scientific Information" has been given in Sweden. It was offered to university students and to graduates with experience in documentation, and was held at the Tekniska Högskolan (Royal Institute of Technology), and was sponsored by the Swedish National Committee on Documentation. Lectures and demonstrations comprised over 100 hours. The purpose was to give a general introduction to documentation focused, however, around information storage and retrieval on computers.

After a general introduction defining "informatology" as a man-machine system for information handling, six main topics were displayed: 1) User's needs, behavior in information retrieval. The readers' habit. The active research worker. Information in an enterprise; 2) Development and growth of scientific information. Sources of information, primary, secondary and tertiary publications. Institutes and organizations in documentation. Organization and costs for information organs within an enterprise. Patents; 3) Media problems: Rules and standards for scientific authorship. Reprography. Translations; 4) Analysis of information: Classification. Coordinated indexing. Abstracting; 5) Experimental methods for information retrieval: Semi-mechanized and computerized systems; 6) Informatological aspects: Codes and error correction. Neural networks. Self-organizing systems. Uncertainty theories, entropy and redundancy. Linguistic problems in information retrieval. Semantic structures in documentary languages.

Of the sixty participants from industries, universities, and libraries, 46 fulfilled the course with high attendance at the lectures. Of the students over 40 were graduated from various universities. However, the difference between the graduates and undergraduates didn't seem to be so cumbersome as did the difference between those with training in science and those without. Therefore, the conceived short introduction into mathematical and logical basis for information transmission had to be expanded. As example of systems in use the Hoffmann-La Roche and Derwent Ringdoc at Karolinska Institute were demonstrated on IBM 1401. The construction of a union catalogue was shown on a Ferranty Mercury by use of paper-tape input.

An evaluation of the course was made by means of a questionnaire to the students at the end of the course. It was shown that about half of the students wanted a better background in mathematics and statistics, that more emphasis should have been given to exercises, that the text-book (C P Bourne, Methods of Information Handling) supplemented by compendia by the teachers was a good choice. For the time being it was obvious that this course had to depend on a team of specialists as teachers. Some overlapping and differences in terminology were, thus, unavoidable.

Plans for the immediate future is to start a similar course in the Fall at Uppsala University. This course will fit into the regular university programme. In order to avoid the introductory training in mathematics, the students missing such training will be requested to attend a special course offered by the statistics department. The course will include about the same amount of lectures, and 60 hours of exercises and special work will be added. The students have to choose between: A) Documentation within the library network, B) Scientific-technical

documentation, C) Managerial and economic documentation, and D) Medical documentation.

For those students who want to proceed in their studies to an intermediate or higher level opportunities will also be open. From July 1965 the Swedish universities will implement a new curriculum in information handling. The earlier disciplines: Numerical analysis, and Data processing will merge under a new umbrella: Information handling. The second term curriculum includes information retrieval and the processing of verbal information. In this way the students will become machine oriented.

References:

P. Atherton, Letter to the Editor. - American Documentation 16,1965, 126.
B. Tell, Curricula for a suggested training course for documentalists. - Library World 64,1963, 346-52.

THE EDUCATION AND TRAINING OF DOCUMENTALISTS IN CANADA. Samuel Rothstein. University of British Columbia School of Librarianship, Vancouver, Canada.

The Canadian pattern of education for librarianship and documentation closely resembles that of the United States. The agencies responsible for such education are graduate library schools. These are departments of universities, analogous in organization and function to other professional schools, such as social work, within those universities. There are now five Canadian library schools, located at the University of British Columbia (Vancouver), the University of Toronto (Toronto), McGill University (Montreal), the University of Ottawa (Ottawa), and the University of Montreal (Montreal). The first three serve the English-language libraries and the University of Montreal the French-language group; the University of Ottawa Library School is a bilingual institution. All five library schools demand a bachelor s degree as prerequisite to admission.

The program of instruction in U.B.C., Toronto, Ottawa and Montreal is of one academic year's duration and leads to the Bachelor of Library Science degree (B.L.S.), which is roughly equivalent to the Master of Library Science (M.L.S.), the first professional degree commonly awarded in the United States. The library science program at McGill has recently been extended to two years and leads to the (sixth-year) Master of Library Science degree. An extension of the University of Montreal program to two years is now contemplated. The University of Toronto also offers a (sixth-year) Master of Library Science degree course and the University of British Columbia will offer a similar course in 1966.

Documentation is generally regarded in Canada as a branch of or as an advanced specialization within librarianship and the training of documentalists is therefore almost entirely a function of the library schools. Most of the library schools offer some orientation or beginning courses in documentation as part of the B.L.S. degree course, but they reserve more intensive and specialized study of this field for the M.L.S. program, where it is likely, indeed, to be a major feature of the course. In addition to such degree programs in documentation, informal workshops and short courses on various aspects of this subject are frequently presented by library schools, library associations and by individual libraries.

The next decade will certainly see the establishment of more library schools in Canada (two are, in fact, already announced as to begin in 1966--at the University of Alberta and at the University of Western Ontario) and large-scale changes in the curriculum. There is considerable pressure to give documentation a greater place in the library school program and it is likely that the present limited offering within the B.L.S. degree course will be expanded. It is also possible that Canadian library schools, in conjunction with computer centres and other interested university departments, may offer degrees in information science. Nevertheless, Canadians will probably retain the view that documentation does not appropriately constitute a field separate from librarianship and that the training of documentalists should therefore be based upon and follow a thorough preparation within librarianship.

Bibliography

1. Bishop, Olga B., "Notes on Library Education As Seen by a Perusal of Accredited Library School Announcements", Institute of Professional Librarians of Ontario Newsletter, 6 (April, 1965), 2-4.

2. Bulletins of the library schools at the University of British Columbia, University of Toronto, McGill University, University of Ottawa, University of Montreal

3. Henderson, Betty, "Library Education in Canada", Canadian Library, 21 (November, 1964), 176-77.

4. Land, R.B., "Library School Curricula" (mimeographed, 8 p.). Issued by the Institute of Professional Librarians of Ontario, 1965.

LA FORMATION DES DOCUMENTALISTES EN FRANCE. Paul Poindron. Paris, France.

Elle est assurée depuis 1950 par l'Institut national des techniques de la documentation (INTD), Conservatoire national des Arts et Métiers. Recrutement direct pour licenciés et ingénieurs. Examen probatoire pour bacheliers. Deux années d'enseignement (72 + 61 heures). 71 diplômés en 1964 sur 125 entrées en 1962. Augmentation croissante des candidats. Insuffisance du nombre des candidats: a) masculins, b) du niveau de la licence, c) ayant une formation scientifique. Collection "Documentation et Information" (Gauthier-Villars) - Pour favoriser la formation technique supérieure dans les Facultés des lettres et des sciences humaines un certificat de technologie (Documentation) a été créé en 1962 à Toulouse et en 1963 à Nancy. Projets pour Lyon - Cours à l'Ecole pratique des Hautes-Etudes sur "Sémiologie et documentation"(J.C. Gardin et B. Jaulin), "Bibliographie et méthodes de documentation dans les sciences sociales" (J. Meyriat) - L'Union française des organismes de documentation (UFOD) assure la formation des aides-documentalistes (2 sessions par an) et organise des sessions (5jours) d'initiation ou de recyclage - Des études en cours (Commission de la recherche scientifique et technique du Ve Plan), à propos des problèmes de documentation de la propriété industrielle, ont porté sur la formation d'un personnel chargé de concevoir et de mettre en oeuvre les méthodes et machines nouvelles nécessaires aux applications automatiques dans le domaine de la documentation. Si les crédits sont accordés, comme on l'espère, il conviendra de déterminer l'établissement ou les établissements d'enseignements compétents, soit organismes nouveaux, soit transformation d'institutions existantes comme l'INTD ou certaines universités.

AREA I. EDUCATION AND TRAINING OF DOCUMENTALISTS

Symposium B: Needs for the Next Ten Years and How to Meet Them

THE PRESENT STATE OF EDUCATION AND TRAINING FOR DOCUMENTATION IN JAPAN AND SOME FUTURE PROBLEMS. Masao Kotani. Faculty of Science, The University of Tokyo, and Yosoji Ito. The University of Tokyo Library, No. 7 Hongo, Bunkyo-ku, Tokyo, Japan.

The status of the regular training courses of library science especially for research or special librarians in Japan is very poor. Although two undergraduate library schools are opened in Japan, the objectives of these schools are to train professional librarians for various types of libraries, so the curriculum is widely generalized and the training for research librarians and special librarians is insufficient. No graduate courses for librarianship and the regular courses for documentalists are founded in Japan. The Japan Library Association, the Japan Documentation Society, and the Ministry of Education frequently open each short term training courses for documentation. Unfortunately, these short term training courses are planned independently by each organization and are not in liaison with each other. It is gradually recognized among scientists of different disciplines, particularly among chemists, that the basic training of literature search, art of writing primary papers, etc., need be included in the curricula of university courses. In this connection the present situation in some universities of Japan will be reported.

INTEGRATED TRAINING POLICY FOR DOCUMENTALISTS ON NATIONAL LEVEL. NEW TRENDS IN THE FEDERAL REPUBLIC OF GERMANY. Dr. Martin Cremer. Institut für Dokumentationswesen, Frankfurt am Main, Germany.

The rapid development of documentation and information in the past ten years is based - apart from organizational and financial demands - on two conditions: effective methods have to be found and applied in order to evaluate the continiously increasing flood of knowledge and to comply with the growing demands of information. Secondly, a sufficient number of specialists have to be trained, who will be charged with these tasks. Both conditions cannot be complied without certain planning activities. With regard to the different educational systems these activities in the field of training may be realized on a national level.

Therefore, in the Federal Republic of Germany the enlargement and integration of training possibilities is prepared. The Deutsche Gesellschaft für Dokumentation has organized until now 8 courses for documentation - each course lasting about 10 months and comprising nearly 400 lessons. This has been a first and successful step, but ist became evident that for present and future needs, these courses are not sufficient. Therefore, the creation of a training centre for Documentation in Frankfurt is planned. This centre will train mainly professional documentalists with or without academic degree. Besides this, special courses for medicine, chemistry, technology, social sciences, sciences and humanities will be arranged, and special fields (indexing, classification, reprography and application of mechanized techniques) will be taught. The centre will also arrange special courses for further education of documentalists giving them the opportunity to inform themselves about the most recent development.

All necessary elements of the science of information and documentation as well as important sections of library sciences with particular regard to the needs of special libraries should be included in the plans for instruction. Thus, a not always existing but desirable integration of documentation and special libraries by common training opportunities is intended.

Therefore, it is to be expected that the number of participants of the different courses will be high, especially the number of participants of the courses for professional documentalists.

The instruction will be done by a small number of full-time teachers and a big number of part-time experts coming from different fields of documentation and library science.

The training centre for documentation should be in close local and technical connexion with a documentation-library, including all instruction fields, and an information centre for literature on documentation, information and library science, with a division of reprography as well for instruction as for practical use, and with the Zentralstelle für maschinelle Dokumentation (ZMD), which will take care of the training on its particular field.

This integrated training system will be completed in a non-centralized way by education of users of documentation at universities, in industry and in administration. Here planning and promoting activities are necessary, too.

EDUCATION FOR SYSTEMS PLANNING. Don R. Swanson. Graduate Library School, University of Chicago, Chicago, Illinois.

The training of documentalists and librarians is usually considered in terms of skills required to operate libraries and document centers. It does not necessarily follow that those who acquire such skills will also be able to competently and imaginatively plan future libraries and document centers; it is education for this latter purpose that constitutes the topic of this paper.

The planning of almost anything involves, first, identification of goals or objectives, and secondly an understanding of the resources for accomplishing those objectives. Resources for accomplishing goals consist of people, machines, information, and techniques of indexing, classification, and other means of "intellectual access" to information. The task of systems planning is one of the optimal allocation of limited resources.

A graduate curriculum addressed primarily to the planning of information systems might reasonably be divided into four categories of courses: (1) The use and users of information, (2) systems planning and computer technology, (3) a study of existing library and information resources, and (4) a study of e indexing, classification, reference and bibliography. Further discussion of these and related categories is given in reference 1.

Undergraduate preparation for such training is a matter to which a great deal more serious thought should be given by educators. There are certain areas of knowledge which I believe provide an especially good foundation for the planning of information systems, and, at least tentatively, would propose the following for consideration:

(1) The language of elementary mathematics and logic.

> The student of systems planning should have good preparation in elementary mathematics at the level of first year college, and to include algebra, simple functions, graphs, and systems of equations and inequalities. Mathematics is here regarded not so much as a specialty but as extension of our powers of communication. Communication in terms of numbers, symbols, graphs, and simple statements of formal logic are prerequisite to the study of the technology of information handling.

(2) Linguistics.

> Syntactic and semantic relationships of words in subject headings and titles constitute an important foundation in any study of indexing and classification. The science of linguistics has been largely detached from studies of indexing but should be regarded as an essential component of educational preparation in this area.

(3) Empirics and statistics.

> Those who plan information systems of the future should be able to understand and interpret the results of research studies in

the information sciences. In many of these studies, data are collected and interpreted. Thus, an understanding of research procedures, methodology, empirical reasoning, and elementary statistics is necessary.

Undergraduate preparation should above all involve the student deeply in the pursuit of knowledge and scholarly inquiry, and in the solving of problems. To this end, the more substantive the curriculum the better, regardless of the area of specialization. In addition to mathematics, logic, and linguistics, those conventional disciplines probably best suited to this type of purpose would include most sciences, especially physics, psychology, econometrics and the more quantitative aspects of the social sciences.

1. Don R. Swanson, ed. The Intellectual Foundations of Library Education. The 29th Annual Conference of the Graduate Library School, July 6-8, 1964. (Chicago, University of Chicago Press.)

FUTURE TASKS IN EDUCATION. Dr. J. Toman. Czechoslovak Academy of Sciences. Prague, Czechoslovakia.

In the field of documentation it is necessary to devote attention to two great problems:

1/ to the education of scientists, technicians and other specialists in the art of documentation

2/ to the accumulation and publication of practical experience of the best methods used during different stages of information processing.

1. The education of specialists in the art of personal documentation.

The "explosion of information" cannot be mastered by the efforts of the information specialist only. Every future scientist, technician and specialist must be educated at school already in the art of personal documentation.

There must come about a revolution in the minds of those who prepare the high school and university curricula. The education at schools in the whole world is based on the assumption that the student must acquire as much knowledge as possible by storing it in his memory. But the capacity of the memory is limited and cannot expand. This problem can be solved only by admitting the fact of limitation of the memory and by teaching the students how to store information in other media than memory. The art of personal documentation must be included in the curricula of schools.

General understanding for the problems of scientific information will be spread only when every student will master the fundamental principles of information technique and when most of them will experience their personal documentation.

We see from experience especially in the pure research that the personal documentation of individual scientists can be well coordinated with the effort of information centres.

The methods of personal documentation must differ from those of information centres. The author of this paper first wrote a manual of documentation technique and soon came to the conclusion that there is a great need for a book about personal documentation. Shortly before his book about the technique of personal documentation /"How to collect knowledge"/ had been published in 1960, he was impressed by the fact that the American scientist Engelbart /Stanford Research Institute, Menlo Park, California/ wrote in the American Documentation about the importance of devoting attention to the elaboration of special methods of personal documentation /"microdocumentation"/ in the same way as the author in Czechoslovakia.

The experience of the author shows that a general book can satisfy a large circle of readers. The scientists only would need a special book containing a higher degree of knowledge, in accordance with their needs.

2. Accumulation of practical experience with different methods of documentation.

The author was entrusted by the Committee FID/TD /Training of Documentalists/ to compare existing manuals of traditional documentation technique. It is interesting to state that most of the newly edited books in the field of information

are devoted to special subjects and to the methods of mechanical retrieval and that there exist only very few books in the world comprising the whole complex of documentation methods, the knowledge of which is so important for the majority of information centres.

The author came to the conviction that we should give our energy to the accumulation of experience with different documentation methods used for different stages of the information process.

It would seem that under these circumstances it is rather difficult to accumulate descriptions of thousands of information systems. But when analyzing many of the information systems we realize that although we do not encounter two similar systems, we find the same methods in them. We come to the conclusion that each information system consists of a number of activities, which can be performed by different methods. As there are on the average no more than 10 different methods used for each of these stages /activities/ we find that the total number of methods, of which the information system can be built does not exceed 100-120 methods.

Thus an information system is only a combination of these different units /methods/. The effectiveness of an information system depends on the choice of suitable methods for different stages of the information process according to the needs and conditions of the organization where the information center works.

The accumulation of experience with all these methods, which form the different information systems, would not demand much time and money and still it would result in a survey of the units from which the suitable information system should be built.

We should devote our attention to the project of this kind. The accumulated experience published in a manual could form the basis of an important subject in the curricula of schools and courses for the training of documentalists.

AREA II. ORGANIZATION OF INFORMATION FOR DOCUMENTATION

Symposium A: Transformation and Organization of Information Content

THE TRANSFORMATION OF SENTENCES FOR INFORMATION RETRIEVAL. Jane J. Robinson. The RAND Corporation, Santa Monica, California, U.S.A.

Sentences occurring in ordinary text are often extremely efficient forms for storing information and a single sentence may be capable of providing answers to many questions. Their surface structures reflect the preoccupation of a writer with his immediate purposes for organizing his data, for emphasizing some aspects of it and subordinating others. Various stylistic choices are also open to him. As a result, semantically equivalent content may be embedded in a great variety of syntactic structures. Moreover, automated parsing grammars assigning structural descriptions directly to sentences in a text encounter great difficulty. If applied heuristically, they may miss a valid structural assignment that correctly correlates an expression with equivalent paraphrases and relevant questions. If applied algorithmically, they tend to produce an unmanageable number of parsings, of which a surprising proportion correspond to possible ambiguities in interpretation and are therefore not eliminable on syntactic grounds.

Recent advances in grammatical theory provide a framework for constructing grammars with a base component to produce or recognize a relatively small set of simple, unambiguous "deep" structures from which the more complex surface structures of sentences are derived by a transformational component which rearranges, conjoins, and embeds. The transformations do not alter meaning, although they are frequently sources of ambiguity. Consequently, the unmanageably large number of multiple analyses produced by loosely constructed grammars applied to surface structures may be reduced by subjecting them to inverse transformations and comparing the remaining structures with a tightly constructed grammar for the simpler deep structures from which any valid surface structures must be derived. Since these deep structures are unambiguous, they can provide appropriate canonical forms for semantic analysis and structural matching in question-answering and deductive systems for information retrieval.

In the last six years, several research groups have attacked the problem of designing automated question-answering systems based on text rather than on highly structured data bases. They have used various techniques of syntactic and semantic analysis, in varying combinations. The view proposed here is that semantic (and other) techniques will prove more effective if they are applied after a syntactic analysis that makes the deep structures explicit.

The major linguistic task is to provide analytic recognition grammars with transformational components adequate to deal with the complexities of a full syntactic analysis, so that the necessarily ad hoc simplifying assumptions of previous question-answering systems can be largely dispensed with. Until quite recently, transformational grammars have been written to generate rather than to analyze and recognize the sentences of a language, although Matthews proposed a technique for analyzing a given sentence by synthesis from a transformational grammar, as early as 1961. Work on the recognition problem has now been undertaken, and three different types of grammar are being provided with transformational components designed to recover deep structures automatically. Kuno reports some experiments with the Harvard predictive analyzer, which are designed to produce kernel sentences con-

currently with the analysis of the surface structure. For phrase structure grammars, three methods have been proposed by S. Petrick, by M. Kay, and by the MITRE Language Processing Techniques Subdepartment. An " approximate" formalism to obtain structural descriptions very similar to deep structures is being developed by Lieberman, et al. at IBM Research. Although applied to a phrase structure grammar, the formalism is intended to be applicable to other models as well. Robinson experimented briefly with a paraphrasing routine for a phrase structure grammar, but is currently designing a dependency grammar with a transformational component, in collaboration with Hays and Kay.

Several machine translation groups are also incorporating transformational features into their grammars, designed to uncover deep structures, in accord with Harris' assumption that many languages are more similar in their kernel sentences than in their total structure. Linguistic work in information retrieval is obviously closely related to work in translation, but no attempt is made here to discuss the latter field.

Bibliography

Fodor, J.A., and Katz, J. (ed.). The Structure of Language: Readings in the Philosophy of Language. Prentice-Hall, Inc., Englewood Cliffs, 1964.

Hays, D.G., and Ma, R. Computational Linguistics: Bibliography. Rm-4523-PR, The RAND Corporation, Santa Monica, California, March, 1965.

Hays, D.G. "Dependency Theory: A Formalism and Some Observations," Language, vol. 40, no. 4, October - December, 1964.

Katz, J., and Postal, P.M. An Integrated Theory of Linguistic Description. Research Monograph No. 26, The M.I.T. Press, Cambridge, Massachusetts, 1964.

Kuno, S. "A System for Transformational Analysis" Preprint, 1965 International Conference on Computational Linguistics. New York, May 1965.

Walker, D.E. (ed.) English Preprocessor Manual. SR-132. Information System Language Studies Number Seven, Language Processing Techniques Subdepartment, Information Sciences Department. The MITRE Corporation, Bedford, Massachusetts, December 1964.

Simmons, R.F. Answering English Questions by Computer: A Survey. SP-1556. System Development Corporation, Santa Monica, California, 2 April 1964.

Salton, G. "Automatic Phrase Matching," Preprint presented to the 1965 International Conference on Computational Linguistics. New York, May 1965.

Robinson, J. The Automatic Recognition of Phrase Structure and Paraphrase RM-4005-PR (abridged), The RAND Corporation, Santa Monica, California, December, 1964.

Petrick, S.R. "A Recognition Procedure for Transformational Grammar," Proceedings of the 2nd Congress in the Information System Sciences. Hot Springs, November, 1964.

Matthews, G.H. "Analysis by Synthesis of Sentences in a Natural Language," 1961 International Conference on Machine Translation and Applied Language Analysis, Her Majesty's Stationery Office, London, 1962.

Lieberman, D., Lochak, D., and Ochel, K. "Automatic Deep Structure Analysis Using an Approximate Formalism", Preprint of a paper presented at the 1965 International Conference on Computational Linguistics. New York, May 1965.

Kuno, S., and Oettinger, A., "Syntactic Structure and Ambiguity of English," AFIPS Conference Proceedings, Vol. 24, 1963 Fall Joint Computer Conference.

SOME LOGICAL ASPECTS OF THEMATIC REPRESENTATION WITH EXPLICIT TIES.
T. M. Williams. Graduate Library School, The University of Chicago, Chicago, Illinois 60637, U. S. A.

Linear sequence, an essential feature of speech transcribed in ordinary writing systems, is a valuable conventional feature in some special notation schemes as well. But when expressions given linearly are listened to or read with some degree of understanding, the linear sequence is in effect attacked, and in general altered: Part of the process of grasping what is said or written consists, in effect, in recognizing which of the sequential features are functional and in distinguishing, for those that are functional, significant differences among the functions locally involved.

Noting this, and noting also the utility of diagrams, tables and index lists, one may propose to dispense with the over-all linearity characteristic of speech and ordinary writing through use of schemes which present or re-present message "content" in another fashion -- a fashion such that explicit ties of several kinds are the basic formal means by which functional components are combined and integrated to form components of greater complexity. By conventions for tie direction, the recorded positions of tied components may then be openly nonfunctional. An abstract model for such a scheme is available in the mathematical notion of graphs (especially, directed graphs) composed of points and (directed) point-connecting lines. Such graphs are most plainly represented by diagrams or three-dimensional constructions or by associated matrices. Straightforward development and plain use of graphic schemes in representing messages or message components can result not only in graphs that are trees but also in graphs that contain cycles and graphs that are nonplanar.

Organic chemists exploit a graphic model quite directly when they choose to represent a chemical compound by a graphic formula or a molecular model, instead of a linear formula or name. For chemists, each of these interrelated modes of representation is a tool available, reputable, and particularly useful on certain kinds of occasion. An analogous latitude of choice could conceivably benefit the work of persons concerned with messaging in a broader way -- for example, persons concerned with information processing in documentation systems, or persons concerned with cognitive aspects of language processing in human minds or brains. Conjectures of what might result, scientifically and technically, from the use of graphic modes of representation in functional connection with linguistic modes seem to invite and justify pursuit of a graphic model in developmental efforts more persistent, penetrating and extensive than those now on record. Such efforts would fall largely within a field called semiotic (the science of signs), which on definition encompasses much of what is specialized in logic and in linguistics.

To consider some matters of logic that are pertinent in a graphic approach to message processing, attention is addressed to a particular project pursuing that approach and to the ways in which such matters are at present treated in it. In the project concerned, themes or topics are represented directly, in a relatively neutral way, and themes play a role in the analysis and representation of propositions and statements. The logical aspects considered are discussed by way of such terms as the following: statements (indicative and other), propositions, and themes (topics); predication and predicates; nonsymmetric predicates; syncategorematic predicates; common names and proper names; connectives; individuals (events

and others); parts of things; quantification; normal forms. The current scheme is exemplified in application to topic-indicating texts.

Bibliography

Note. The following bibliography is confined to several introductory works (one each for linguistics, graph theory and logic), one seminal work that is easily overlooked (Kempe), and one paper by the present author containing references to a few of many further works that might be cited.

1. Gleason, H. A., An introduction to descriptive linguistics, d 1961 (revised edition), Holt, Rinehart and Winston, Inc.

2. Kempe, A. B., A memoir on the theory of mathematical form, d 1886, Philosophical transactions of the Royal Society of London v 177 part 1 p 1-70. (For an abstract and corrections, see Proceedings of the Royal Society of London d 1885 v 38 p 393-401; d 1887 v 42 p 193-196.)

3. Ore, O., Graphs and their uses, d 1963, Random House.

4. Quine, W. V., Elementary logic, d 1965 (revised edition), Harper & Row, Publishers.

5. Williams, T. M., Topic charting and paraphrase collecting through use of a multidimensional grammar, d 1963, p 47-50 in H. P. Luhn (ed.), Automation and scientific communication, short papers, American Documentation Institute.

THE TRANSFORMATION AND ORGANIZATION OF INFORMATION CONTENT: CONTRIBUTION OF PSYCHOLOGY. Philip J. Stone Dept. of Social Relations, Harvard University Cambridge, Massachusetts, U.S.A.

Since psychologists work with information theorists and documentation specialists at a number of junctures, it often becomes somewhat arbitrary to determine what ideas are to be credited to what disciplines. Contributions begun by specialists in one field may be developed by specialists in another. Such organizations as the RAND or System Development Corporations in the United States or the EURATOM or C.N.R.S. projects in Europe have had a particular significance for interdisciplinary work in this area.

Without being too presumptuous in presenting claims for my field, certain viewpoints of a distinct psychological bearing can be presented as perhaps having a significant future role in documentation research:

Organization of information

1) View of man as a limited "list processing" organism.
   -limitations in lengths of lists.
   -limitations in organizations of lists.
   -limitations in handling disjunctive concepts.
   -tendency to ignore negative information, i.e. to ignore considering subsets of items that all share in not having a certain characteristic.

2) Use of tree structures to represent human list processing.
   -types of tree structures
   -Yngve's "seven plus or minus two" list length hypothesis taken from G.A. Miller
   -Feigenbaum's "EPAM" representation of nonsense syllable learning.
   -Hunt's concept learning model.
   -alternative languages for representing list structures on a computer: IPL-V, LISP, COMIT, SNOBOL, DYSTAL, SLIP, etc.

3) Development of documentation techniques to aid in extending human list processing capabilities; emphasis on "shuffling" information according to presence or absence of characteristics.
   -Hunt's model: emphasis of logical characteristics
   -Sonquist's model: emphasis on statistical characteristics
   -representing shuffling of tree structures on "Project Mac" time sharing system, using typewriter console.
   -example results.

Units of Organization: Defining Basic Elements

1) Humans tend to "chunk" information at a rather high level of abstraction.
   -fashions of jargon often represent different chunking patterns.
   -after chunking takes place, composite elements tend to be forgotten; chunks tend to be taught as unanalyzed entities.

2) Within a field,"building block" elements can be identified.
   -"extension" referents: basic measurement procedures used in research.
   -"intension" properties: basic characteristics considered to be relevant to field.
   -implications of identifying further substructures.

- examples of "building block" identification in field of survey research.

3) Importance of identifying "building block" elements within newly developing fields.
-should be a responsibility of the profession, not left to librarians and later historians.
-"building block" elements provide a flexible cognitive basis for structuring and comparing different theoretical fabrics; places the biases of an era into an empirical and theoretical perspective.

Bibliography

Carnap, R. Meaning and Necessity, Chicago: Phoenix Books, 1947.

Feigenbaum, Edward A. Computers and Thought, New York: McGraw-Hill, 1963.

Hunt, E.B. Concept Learning, New York: Wiley Press, 1962.

Hunt, E.B., Kreuter, J., and Stone, P. Experiments in Induction, New York: Academic Press, in press.

Licklider, J. C. Concepts and Problems of Libraries of the Future, Cambridge: M.I.T. Press, 1964.

Miller, George A. "The magical number seven plus or minus two, Psychological Review , 1956, 63, 81-97.

Scheuch, Erwin K. and Stone, P.J. "The general inquirer approach to an international retrieval system for survey archives." The American Behavioral Scientist.Vol. 7, No. 10. June 1964. pp. 23-28

Sonquist, John A. and Morgan, James N. "Problems in the analysis of survey data, and a proposal", Journal of American Statistical Association, 58 (June 1963) pp. 415-435.

Yngve, V. "A model and an hypothesis for language structure" Proceedings of the American Philosophical Society. Vol. 104, No. 5 October 1960. pp. 444-466.

TRANSFORMATION AND ORGANIZATION OF INFORMATION CONTENT: CLASSIFICATION RESEARCH. Phyllis A. Richmond. University of Rochester Library, Rochester, New York, U. S. A.

Recent developments in classification research are discussed in terms of philosophic approach and methodology. A series of definitions is given in order to clarify the discussion and to indicate parameters of the study. The relationships of logic, mathematics, scientific and probability methods, and linguistics are explored within the framework of current research in classification. Three major types of research procedures are distinguished, and the various schools of thought involved are tentatively identified. The necessary basic intellectual problems to be solved are singled out for attention and the future of classification research efforts outlined with regard to these problems.

BIBLIOGRAPHY

1. Alston, William P. Philosophy of Language. Englewood Cliffs, N. J., Prentice-Hall, 1964.

2. Barker, Stephen F. Philosophy of Mathematics. Englewood Cliffs, N. J., Prentice-Hall, 1964.

3. Black, Max. Critical Thinking; an Introduction to Logic and Scientific Method. 2nd edition. New York, Prentice-Hall, 1952.

4. Cherry, Colin. On Human Communication, a Review, a Survey, a Criticism. New York, Wiley - Science Editions, 1961.

5. Chomsky, Noam, "Logical structures in language," American Documentation, V. 8, No. 4 (Oct. 1957), pp. 284-291.

6. Fairthorne, Robert A. Towards Information Retrieval. London, Butterworths, 1961.

7. Ranganathan, S. R. Prolegomena to Library Classification. 2nd edition. London, The Library Association, 1957.

8. Salmon, Wesley C. Logic. Englewood Cliffs, N. J., Prentice-Hall, 1963.

AREA II. ORGANIZATION OF INFORMATION FOR DOCUMENTATION

Symposium B: Comparison and Evaluation of Transformation Techniques and Organizational Structures in Documentation Systems

THE EFFECT OF A CITATION INDEX ON LITERATURE USE BY PHYSICISTS. Ben-Ami Lipetz. Carlisle, Massachusetts, U. S. A.

An experiment was conducted for the American Institute of Physics Documentation Research Project for the purpose of evaluating the impact of a citation index to physics literature upon the literature-use habits of physicists. A citation index was specially prepared in which it is possible to look up a known reference in either of two very heavily used United States physics journals -- The Physical Review and the Journal of Applied Physics -- to learn whether, and precisely where, that reference was cited in four leading Soviet physics journals which are published in English translation but which have only about one tenth as many subscribers in the United States. The Soviet journals which were used as citation sources in the index were Soviet Physics - Crystallography, Soviet Physics - JETP, Soviet Physics - Solid State, and Soviet Physics - Technical Physics. It was hypothesized that because the citation index would help United States physicists to learn of Soviet papers with subject contents similar to the known United States references they would start with, it would therefore tend to increase the utilization of the Soviet physics translation journals by United States physicists.

The experimental citation index was prepared by punched-card techniques. The Soviet source journals which it covered dated from 1955. About 18,000 citations of the two United States journals were found, mostly citations of The Physicial Review. The citations were rearranged for printing into year, volume, and page order for each of the cited journals.

The impact of the experimental citation index was measured principally by means of statistics which were collected monthly on utilization of the Soviet physics translation journals in libraries. (There are almost no individuals subscribing to these translation journals.) More than one hundred libraries throughout the United States contributed data to the study. The statistics were in several categories, reflecting borrowing activity, photocopying activity, and display-issue use. Statistics were reported separately for each of the four Soviet physics translation journals used as sources in the citation index, and for an additional four Soviet physics translation journals which were excluded from the index to serve as controls. Statistics were supplied by the participating libraries for a period of months preceding and following the controlled distribution of the experimental citation index. The index was distributed without advance notice to the 550 subscribers to The Physical Review and/or the Journal of Applied Physics in a restricted, self-contained geographic region which accounts for five percent of the domestic subscriptions to American Institute of Physics publications. Some 85 percent of these subscribers are individual physicists, rather than libraries. Library-use statistics from participating libraries in the test region were compared with statistics from participating libraries in the remainder of the country in order to evaluate the impact of the citation index.

The study of library use was supplemented by a parallel study of changes in the subscriber lists for the test journals as compared to the control journals following distribution of the citation index, and by a study of back number orders for the source journals and control journals received by the publisher before and after

distribution of the index. Each recipient of the citation index was contacted by mail to determine whether he had tried to use the index and to what extent. Personal interviews were conducted with a sampling of the index recipients.

The impact of the citation index on literature use by physicists -- as indicated by changes in the ratio of test-region library statistics to control-region library statistics following distribution of the index -- was anything but dramatic. Although a definite impact was detected in the first six months or so after distribution of the index, it averaged out to only about a 15 to 20 percent increase in borrowing and photocopying and about a 40 percent increase in display-issue browsing; at the end of this period the utilization figures appeared to have returned to normal levels a few percentage points higher than before the index was distributed. Since utilization levels were very low to start with, the impact of the citation index in terms of actual stimulated literature utilization is very small. The statistics gathered in this experiment permit the rough calculation of the number of scientist-hours which were diverted to the source literature by the experimental index.

Results obtained from the other evaluation techniques used in this experiment are in good agreement with the library-use statistics. No indications of a large impact attributable to the citation index were found in the statistics on source journal subscriptions or in the statistics on back-number orders which were supplied by the publisher. A survey questionnaire which was sent to all citation index recipients a few months after distribution of the index confirmed the fact that only a minor fraction (42 out of the 260 who replied) had tried to use the index for information retrieval.

Despite their low usage of the experimental citation index, the comments of physicists who had received the index indicated a generally favorable attitude toward the citation concept in literature retrieval. Those physicists who had actually used the index tended to be extremely enthusiastic in their comments.

It is concluded that although the citation index concept is by no means a cure-all for the literature problems of physicists, it can nevertheless be of definite value as a supplement to the existing "universe" of reference aids. It deserves further study and development, and should be taken seriously. No conclusions can be drawn from this experiment regarding the effectiveness of the citation index as substitute for any specific conventional reference aid or aids.

1. B. A. Lipetz, "Design of an experiment for evaluation of the citation index as a reference aid," pp. 265-266 in Automation and Scientific Communication, Preprints of the Annual Meeting of the American Documentation Institute, Chicago, October 6-11, 1963.

2. B. A. Lipetz, assisted by C. F. Dziurkiewicz, Citation of Physics Literature. Experimental Issue, New York: American Institute of Physics, 1964, 105 pp. (AIP/DRP-C1 1964).

AUDITING PROCEDURES FOR INFORMATION RETRIEVAL SYSTEMS. Björn V. Tell. AB Atomenergi, Nyköping, Sweden.

The need for effective information retrieval systems has been most strongly felt in science and technology. The researcher in the information field seems to have been influenced to apply models developed in some of these science fields, e.g. information theory, automata theory, switching theory etc. Information systems both of nationwide "macro"-size and down to the "micro"-size of that of a documentation center in an enterprise involve the use of manpower. In many cases we have to deal with man-machine systems. Instead of looking for an applicable model from the science field, it is here suggested that good use can be made of models from the social sciences. Economic models display many features of interest for analysis both of macro- and micro-information systems.

For a long time the pragmatic aspect of documentation has been stressed by raising the question about relevance in information retrieval. But anyone who stresses "relevance" as a criterion for evaluation purpose, must answer the question - relevant to whom? and for what purpose? Thus, the documentalist is trapped in the same situation as the economist when he discusses "utility" or "value" as an important parameter. However, the economists have developed important methods without settling upon this question. This paper concentrates on dealing with a standard costs model for evaluation purposes, leaving the question about the very nature of relevance open.

Like an auditor the documentalist who is going to perform an analysis of an information retrieval system, has to test his results against some standard of which he is aware. Standards are scarce in the information field due to the fact that most systems often are in a state of transition. The "Proposed Standard Description for Reporting Evaluation Tests of Retrieval Systems" represents a concensus about pertinent parameters for describing an IR-system. The quantitative data in this standard form deal with items, such as documents, index terms, time, costs, personnel and equipment. From an auditor's point of view, money is the common denominator in terms of which the exchangeability of these items can be measured.

There is a market price for many of the items in the standard form, but for some of the parameters we have to make use of estimated costs. As in accountancy, if we use estimated costs, the results will be less accurate than based on market prices, but they might well prove useful for the analysis. The items which have no market price can be converted to expenses by adopting some conventions for conversion.

An auditor's evaluation is concentrated on a study of the results. However, these results presuppose a forecast or a plan according to some goal, and his evaluation consists of a comparison between a plan and its execution. The analysis of variations from standard costs is a review technique which operates by means of a comparison between the results expected according to the plan and the results as revealed in the cost accounting. The results expected according to the plan is designated by the term "standard".

The goal of an IR-system includes the maximization or minimization of one or more factors. However, the present performance of a system can also be regarded as

a goal or a plan. If during a stated time period, the observed data are taken as the intended goal for that period, this will fix a starting point or a "standard" for further planning. Suggestions for changing the plan in order to maximize one factor, will result in a new plan, the execution of which will take place in the new time period. The results of the plan on completion can be evaluated by a study of the variations from the standard.

Attention is called to the irrelevant documents which are treated as "scrap", i.e. the cost for their retrieval has to be charged against the value of the retrieved documents. That the "scrap" can be reprocessed for new questions is accounted for.

Examining the standard costs variations opens a possible way to explain differences in performance within a system at different times. These explanations can be used to formulate changes in the initial conditions in such terms that the changed or new model can serve as an explanation of the observations which led to the rejection of the old one. It is suggested, finally, that changes should be undertaken as long as the costs of the changes are less than the increased value of the system.

References.

Bourne, C.P., Review of the criteria and techniques used or suggested for the evaluation of reference retrieval systems. (Working paper for the NSF Conference on the evaluation of reference retrieval systems and procedures. 1964).

Bourne, C.P. & Ford, D.F., Cost analysis and simulation procedures for the evaluation of large information systems. - Am. Doc. 15, 1964, 149-9.

EVALUATION PROCEDURES FOR COMPUTER-BASED RETRIEVAL SYSTEMS, Gerard Salton. Computation Laboratory, Harvard University, Cambridge 38, Massachusetts, U.S.A.

The interest in evaluation procedures for information retrieval systems and techniques stems from two principal causes: first, more and more retrieval systems are being designed, thus raising a question concerning performance and efficacy of these systems; and second, evaluation methods are of interest in themselves, in that they lead to many useful problems in test design and performance, and in the interpretation of test results.

The work reported here differs from other efforts on systems evaluation in that it deals with the evaluation of automatic, rather than conventional information retrieval. The computer-based information systems of the future may be characterized by the following principal properties:

a) the systems will probably operate in a time-sharing environment, in such a way that a multiplicity of users may be given simultaneous access to the files;

b) the search function may be undertaken iteratively, possibly under user control, by performing several partial searches to approach the desired subject area, rather than a single one-shot process;

c) several different analysis and search procedures may be incorporated in such a system, including, in particular, stored intellectual aids for vocabulary normalization;

d) the size of the document file to be searched may be expected to consist of 100,000 or more documents; under the circumstances, a manual assessment of relevance of all documents with respect to each search request will become impractical, and sampling techniques will be needed to measure retrieval effectiveness;

e) the computation of sophisticated correlation coefficients between documents and search requests will make it possible to present ranked document output in answer to the search requests, in decreasing order of the correlations between documents and request.

In the present report, evaluation techniques and procedures are discussed which are specifically based on the foregoing systems organization. Such important systems parameters as cost of retrieval, response time, influence of physical layout, personnel problems, and so on, are disregarded, and the emphasis is instead on the evaluation of retrieval (that is, analysis and search) techniques.

Since the evaluation process is restricted to automatic systems, a number of human problems which complicate matters in a conventional evaluation situation, including, for example, the difficulties due to inconsistency among indexers, or to the presence of search errors, are no longer so crucial. The evaluation process can therefore concentrate on the main systems parameter: its ability to locate relevant information, while rejecting nonrelevant material.

The design of automatic evaluation systems is first discussed in detail. Recall-like and precision-like measures are then introduced which seem appropriate

for the evaluation of computer-based retrieval systems. The use of these measures in evaluating retrieval performance of automatic, iterative, search systems is discussed, and test results are presented showing results obtained by using the fully automatic SMART retrieval system, now operating experimentally on the IBM 7094. (This system incorporates a large number of automatic search and analysis procedures, and makes it possible to evaluate the effectiveness of each one by processing the same search request against the same document collection in many different ways, while comparing results in each case.)

The prospects and problems of automatic search systems are summarized, with reference in part to existing systems, such as the NASA, MEDLARS, and DDC systems, and tentative conclusions are drawn concerning the design of fully automatic information systems.

REFERENCES

1. Salton, G. "A Document Retrieval System for Man-Machine Interaction," Proceedings of the 19th ACM Annual Conference, Philadelphia (1964).

2. Salton, G., and M. E. Lesk. "The SMART Automatic Document Retrieval System - An Illustration," Communications of the ACM, Vol. 8, No. 6 (June 1964).

3. Salton, G. et al. Information Storage and Retrieval, Reports No. ISR-7 and ISR-8 to the National Science Foundation. The Computation Laboratory of Harvard University (June 1964 and December 1964).

4. Swets, J. A. Information Retrieval Systems Science, Vol. 141, No. 3577 (19 July, 1963).

5. Cleverdon, C. W. "The Testing of Index Language Devices," ASLIB Proceedings, Vol. 15, No. 4 (April 1963).

6. Cleverdon, C. W. "The Testing and Evaluation of the Operating Efficiency of the Intellectual Stages of Information Retrieval Systems," International Study Conference on Classification Research, Elsinore (September 1964).

7. Swanson, D. R. "Searching Natural Text by Computer," Science, Vol. 132, No. 3434 (21 October, 1960).

8. Swanson, D. R., "The Evidence Underlying the Cranfield Results," The Library Quarterly, Vol. 35, No. 1 (January 1965).

9. Rocchio, J. "Performance Indices for Document Retrieval Systems," Information Storage and Retrieval, Report No. ISR-8 to the National Science Foundation. The Computation Laboratory of Harvard University (December 1964).

10. Fairthorne, R. A. "Basic Parameters of Retrieval Tests," short paper presented at the 1964 ADI Annual Meeting, Philadelphia (October 1964).

11. Salton, G. "The Evaluation of Automatic Retrieval Procedures - Selected Test Results Using the SMART System," Report No. ISR-8 to the National Science Foundation, Harvard Computation Laboratory (December 1964).

12. Henderson, M. M. "Tentative Bibliography on Evaluation of Information Systems," National Bureau of Standards Report (February 1965).

DOCUMENT RETRIEVAL SYSTEM EVALUATION PRINCIPLES. Vincent E. Giuliano. Arthur D. Little, Inc., Cambridge, Massachusetts, U. S. A.

The discussion is focused on two main themes: identification of significant measurable characteristics of retrieval system performance, and development of adequate mathematical models to serve as a basis for making measurements. The performance characteristics to be measured must be clearly relatable to the operationable objectives of a specific retrieval system at hand, and must take into account appropriate boundary conditions such as nature of collection, user population, and expected kinds of search requirements. Quite different mathematical models and measuring techniques might be appropriate, depending on the objectives and boundary conditions at hand. Even within a given set of boundary conditions, the mathematical model should provide means for taking into account the numerous sources of variance encountered in practice--particularly, variance in depth-of-search requirements among different users, variance in relevance judgements among system evaluators, and variance in system behavior from one query to another.

AREA III. INFORMATION NEEDS OF SCIENCE AND TECHNOLOGY

Symposium A: Methodology

INFORMATION NEEDS OF SCIENCE AND TECHNOLOGY - BACKGROUND REVIEW. Gerald Jahoda. Florida State University, Tallahassee, Florida, U.S.A.

Studies of information gathering habits of scientists and engineers (SIG) are conducted to learn how and why scientists and engineers use information and information systems (libraries, technical information centers, personal indexes, and other sources of information). These studies also reveal scientists' and engineers' opinions of existing and proposed information systems. The results of these studies are intended to provide information specialists (librarians, information scientists, etc.) with a set of guide lines that will help in the evaluation of existing information systems and in the design of improved information systems. Despite the fact that numerous SIG have been conducted in recent years, basic questions, such as, what is a good information system?, still remain unanswered.

A composite picture of scientists' and engineers' information needs, uses and sources is drawn in this paper from reports of SIG. SIG techniques are examined in terms of type of information that each technique can provide, reliability, and other characteristics.

Computer-based information systems, as exemplified by selective dissemination of information systems and the Project MAC Index at M.I.T., open up new techniques for collecting records of information use as a by-product of system use. Records of use of computer-based information systems can be obtained without disturbing the user and without the need to rely on memory or other subjective factors.

Different aspects of the problem of information needs and uses have to be studied with different SIG techniques. Some aspects of the problem, as exemplified by the use of indexes by future generation of scientists, cannot be determined by any technique. While SIG techniques are far from perfect and leave some questions either unanswered or incompletely answered, they are a basic tool in the design and evaluation of information systems.

USER'S PERCEPTION OF EDITORIAL CONTENT. Joseph H. Kuney. American Chemical Society, Washington, D. C., U. S. A.

In a personal interview survey designed to guide future editorial decisions of a scientific journal, an attempt was made to collect information on the user's reading environment, reading behavior, attitudes and opinions of the journal, and respondent's perception of new editorial subjects. In addition, data on respondent's age, educational background, and publishing pattern were collected.

Personal interviewing of users in an effort to determine their information needs requires careful attention to determination of study objectives, selection of sample, collection and tabulation of data, and finally interpretation of the data.

The objectives of a survey are not fulfilled until the results have been analyzed, interpreted, and then applied. Frequently the latter is the most difficult step to perform since the replies are rarely clear-cut guides to action. When the survey is used as a continuing aid to editorial judgment, it serves as a valuable tool to journals seeking to meet the information needs of scientists.

AN EVALUATION OF THE METHODOLOGY OF THE DOD USER NEEDS STUDY.

Lawrence Berul and Allan Karson. AUERBACH Corporation, Philadelphia, Pa., U.S.A.

User-needs studies can be put into two major classifications: those of narrow scope for design or improvement of a specific information system and those of broad scope for the planning or improvement of a network of information systems. Most of the user-needs studies conducted to date belong in the first category. They have been confined to either a narrow segment of the technical community, e.g., the users of a particular library, or to a narrow area of investigation, e.g., the media used to acquire information. The study recently completed by DOD is one of the few user-needs studies conducted that fits into the second category, and it certainly is the largest.

The DOD study was extremely broad in scope, both in terms of the interdisciplinary nature of the population surveyed and of the areas of investigation. The principal purpose of the DOD study was to collect and analyze a statistically significant data base on how DOD scientists and engineers presently acquire and utilize technical information in the performance of their tasks. The population surveyed comprise the entire research, development, test, and evaluation (RDT&E) community within the Department of Defense. Its 36,000 members represent practically every technical discipline and are engaged in a wide variety of interdisciplinary tasks of a scientific, engineering, or administrative nature.

The methodologies generally used in the previous user studies of narrow scope were not appropriate to a study of the broad scope and magnitude of the DOD User Need Study. Consequently, it was necessary to utilize some survey techniques which had never been previously applied to information needs studies. For example, a principal feature of the survey methodology was the application of the critical-incident interviewing technique. This involved the identification of a task recently completed by the respondent, and the isolation and description of the actual information utilized in the performance of that task.

The critical-incident technique prevented the study from becoming an opinion poll and thereby eliminated the primary source of bias in most user studies. A semi-structured interview guide and handbook were employed to aid the interview process. The structured portion increased the consistency of the question interpretation and simplified the recording of the responses. The unstructured portion was designed to allow for the identification of unanticipated information patterns.

The methodology best suited to a particular user needs study depends upon the objectives of the study. The methodology of the DOD study is not particularly applicable to system-requirement studies confined to a single discipline or a relatively small organization. In these types of studies, it usually is practical to use a considerably less structured survey technique, which facilitates deeper probing of promising areas as they develop, rather than restricting the interview to a rigid format. It may also be fruitful in such cases to determine the information requirements of specific projects or organization functions rather than a random sampling of recently completed tasks.

On the other hand, the methodology used in the DOD study will be highly applicable to conducting future user studies of large heterogeneous populations. The methodology provides for relating specific information requirements to specific task descriptions. This, therefore, produces reliable data useful in the overall planning of a comprehensive network of information systems before attempting to design or improve various systems and subsystems which only service the needs of individual segments of the population.

AUERBACH Corporation

"DOD user needs study. Phase 1, in two volumes."
AUERBACH Corp., Final Technical Report 1151-TR-3, submitted to
Advanced Research Projects Agency, Department of Defense, Washington, D.C.
484 p., May 14, 1965

AREA III. INFORMATION NEEDS OF SCIENCE AND TECHNOLOGY

Symposium B: Increasing the Efficiency of Information

INCREASING THE EFFICIENCY OF THE USE OF INFORMATION - A BACKGROUND REVIEW.
Ezra Glaser. U. S. Patents Office, Washington, D. C., U. S. A.

About Scientific and Technological Information

There is great diversity in the format and content of scientific and technological information. For illustration, a classification of uses is offered which covers the range from specific, well defined data to speculative search of a general problem area.

Increased efficiency might be sought by classifying users (bench chemist, research administrator for a pharmaceutical laboratory) and kinds of information and investigating their interactions.

About Combinations of Elements of Information

Some work has been done on frequency of "association" among ideas and on plausible associations in one-step and higher order links. Such approaches can utilize the abilities of computers and might be used to extend the inquiry, to make for more fruitful search, or to stimulate thought. The particular result depends upon system characteristics not identified here.

There is also a case for unlikely combination - for stimulative effect. This possibility has not received much attention. There is some evidence that unlikely combinations and the unconventional application of known techniques or apparatus promote new directions in creative thought.

Increased efficiency might be sought by abandonment of the one-document or one-piece-of-information approach for much of the user spectrum and pursuing research on combinations and linkages. This approach focuses on the information itself rather than the document or index term.

About People

Individuals exhibit wide ranges of creative talent, knowledge, needs. Each person in each situation has his present compromises - unstated assumptions about what it is futile to ask for - and these are certainly far from the ideal satisfaction of all useful needs for access to recorded knowledge. His statements about his needs will almost always reflect his bondage to present habits of obtaining and using technical information.

In practice, his statement of need prior to search will frequently be inconsistent with his success rating after the search - for the later classifications suggested above - because he learned something during the search. /1/ There is even some evidence that the modifications of statements about need are different for manual and mechanized search /1/.

Increased efficiency of use might follow more studies of real need, as distinct from conventionalized and compromised statements of need.

About Language and its Representation

Scientists communicate in coded languages as well as natural languages; chemical formulas, mathematical formulas, circuit diagrams, mechanical drawings, graphs, block diagrams, maps, photomicrographs, etc.

In such specialized languages, they think, query, make statements, record information, and screen collections of information for their own purposes.

Some of the coded forms of scientific information has been shown to be receivable by computers, thus opening the way to a variety of useful mechanized logical processes. The early experience appears useful but there is a paucity of disciplined experimentation.

Increased efficiency seems likely to follow progress in the technics of computerized manipulation of information, provided that the specific logical operations can be imbedded in appropriate search and retrieval systems.

About Tests and Evaluations

Formal evaluation of actual search experience poses problems of classic difficulty: the activity is largely mental and covert; the subjects are knowledgeable and clever human beings; the object of the activity is subject to change in midstream; and scoring poses difficult and subjective problems. Yet progress depends as much upon valid testing as upon technical achievement.

About Technics and Organization

The above discussion relates to ways of doing things. There is a whole range of additional factors that should affect our efficiency in the use of information - what might be called organizational matters. They relate to the assignments of functions and responsibilities, the provision of financial and technical support, the making of choices of strategy and program.

They are left for one of our distinguished discussants, Dr. Alvin M. Weinberg, whose recent report on these matters has had unprecedented influence. /2/

References: /1/ June R. Cornog, "The Patterns of Thinking in Searching Patent Applications by Manual and Machine-Assisted Methods", In press: Journal of Chemical Documentation. /2/ Government, Science, and Information, A report of the President's Science Advisory Committee, Panel on Science Information, Alvin M. Weinberg, Chairman, January 10, 1963.

THE M. I. T. TECHNICAL INFORMATION PROJECT. M. M. Kessler. The Libraries, Massachusetts Institute of Technology, Cambridge, Massachusetts, U. S. A.

The Technical Information Project (TIP) at the Massachusetts Institute of Technology is an experiment in information-system design. It provides a facility to investigate by direct experience what contributions modern technology can make in solving the problems of scientific communication. The system will be described and demonstrated at the FID Meeting by teletype connection between the lecture hall in Washington, D. C. and the installation in Cambridge, Massachusetts.

The system consists of five major components: a sample literature, a computer facility, a library of programs, a population of users, and a test and monitor procedure.

Elements of the incoming literature are key punched, edited, compressed, and transferred to magnetic tape. This tape is kept as a back-up record. For processing purposes the material is transferred from tape to the disc memory where it is always available for use.

The material in store is taken from 25 journals of physics. For each of the articles in each of these journals we record the location of the article (journal, volume, page), the title, authors, the institutional affiliation, the citations (journal, volume, page), and the location of the article in Physics Abstracts. Periodic monitoring programs test the integrity of the data.

The computer facility is itself an experiment (Project MAC). It consists of a central machine with 150 remote consoles. The consoles are standard teletype units distributed largely around the Massachusetts Institute of Technology. Contact with the computer is by means of telephone. The 150 consoles are available to perhaps 500 people who may at any time use the computer on a time sharing basis. Thirty people may use the computer at the same time.

The computer operates on the literature through a set of commands such as "Search Physical Review Volume 128 and print the titles of all papers by John Smith". Searches may be initiated by author, title, citation, index and/or bibliographic coupling. The search range and the output format commands are flexible and under user control. Several logical tools are available to the user such as "and", "or", "but not". The communication facilities of the system may be used to transmit messages from system to users, from users to system and between users. The user may also receive a print out of all available programs. This facility is being elaborated into a teaching program so that new users may be instructed by the computer.

The use of the system as a research tool will be demonstrated at the Meeting.

CASE STUDY: INCREASING THE EFFICIENCY OF THE USE OF INFORMATION, AEROSPACE RESEARCH APPLICATIONS CENTER. Dr. Howard L. Timms. Aerospace Research Applications Center, Indiana University, Bloomington, Indiana, U. S. A.

During the past two and one-half years, the Aerospace Research Applications Center (ARAC) at Indiana University has been providing information retrieval service of various types to the scientists, engineers, and technicians at subscribing companies totaling forty-five at the time this abstract was prepared. The ARAC project is the first university-based information center operating under contract with the National Aeronautics and Space Administration (NASA). The project's objective is to develop methods of disseminating to private companies the scientific and technical information generated in the nation's space programs. It is believed that this information may be useful in the development of new and improved products, processes, and materials by private companies for nonspace and nonmilitary markets. The project is one of several being conducted by NASA under a requirement of the 1958 Act of Congress that established NASA.

The major information services provided by ARAC are as follows:) retrospective searching of the NASA file now totaling nearly 150,000 documents, (2) selectively searching incoming documents against company interest profiles to provide a current awareness service to company "interest centers." Incoming documents total nearly 4,000 per month. Selective dissemination searches against interest profiles are made twice-monthly. Other services are provided by the Center. They are described in the paper that is abstracted here.

The NASA document file is prepared by Documentation, Inc., NASA's subcontractor that processes raw documents resulting in the file input to ARAC. This file is received monthly by ARAC in two forms: magnetic computer tape and microfiche. Also received are the corresponding abstracts in microcard form.

Increasing the efficiency of the use of this information may be treated in two parts: the first concerns organization on the part of the subscribing companies for efficient ingestion of scientific and technical information; the second part concerns preparation by ARAC of its output to subscribing company people.

Concerning organization on the part of subscribing member companies, this subject is extremely complex and few principles of an operative nature are as yet available to guide companies in this effort. The results of formal efforts by ARAC and company representatives to establish such principles are described in the paper abstracted here. These formal efforts were in the form of a workday on one occasion, and a panel discussion a year later. Over its period of operation to date ARAC has worked informally with its member companies on specific organization for information transfer at each company. The results of these efforts, which for the most part have been successful, are described in the paper abstracted here. They involve specific communication and coordination procedures between ARAC operating personnel, company information users, and company information centers and libraries.

Concerning the preparation of its output for efficient use by company people, ARAC early established several guiding policies that appear to have served very well to date. A few of these policies, which are treated fully in the paper abstracted here, are as follows: (1) design the output in such a manner as to fully

suit the user's needs, not in a manner that one might think the user ought to accept it--in other words the user's needs and desires rule, as in any well-designed consumer-oriented product; (2) never send the user a full-copy document until he has seen its abstract; (3) don't badger the user with questionnaires, forms, and other post-operations control devices--build only the minimum necessary, simply-designed, feedback and control features into the operational paperwork.

Weimer, Arthur M., and Timms, Howard L., "Case Report: The Aerospace Research Applications Center," Business Horizons, Vol. VII, No. 2 (Summer, 1964).

Harvey, Roger K., "The Aerospace Research Applications Center: Programs and Progress," Indiana Business Review, Vol. 40 (April, 1965).

AREA IV. INFORMATION NEEDS OF SOCIETY

Symposium A

CAPABILITIES OF INFORMATION SYSTEMS VERSUS NEEDS FOR INFORMATION.

J. C. R. Licklider. Thomas J. Watson Research Center, International Business Machines Corporation Yorktown New York, U.S.A.

The aim of this paper is to relate capabilities and limitations of present information systems to the needs of society for information. The needs have been set forth in previous papers of this section of the Congress. "Present information systems" subsumes both communication systems such as the mass media, mail, telephone, and telegraph and storage-communication-and-processing systems such as command and control systems, management information systems, document systems, and libraries. The latter set, however, constitutes the focus of this discussion. The capabilities of those systems are examined in respect of storage, communication, and processing capacities and of their abilities to meet the needs for pertinence, timeliness, convenience, and economy.

A main conclusion is that, except in one area, the capabilities of the hardware components of present information systems promise more than the over-all systems deliver, that the bottlenecks of information technology lie not in component devices but in the area called "software," in system planning and design, and in the overcoming of organizational and social intransigence and inertia. The exception is the interface between information machines and people, a domain in which device technology has lagged, but in which there is now much interest and good prospect for advance.

POLICY PROBLEMS OF A DATA-RICH CIVILIZATION. Harold D. Lasswell. Yale University, New Haven, Connecticut, U. S. A.

Open access to sources of information and freedom of dissemination will be restricted so long as the arena of world politics is constrained by security considerations. At present world public order provides an incomplete system of national security. Hence every nation state relies on covert as well as overt information.

If information becomes more inclusive and realistic, political decisions may become more focused and rapid. If the United Nations or any organization of states is permitted to improve its sources of information, security estimates may become more reliable for everyone. Hence the possibility of surprise at least can be largely eliminated.

Allowance must be made for the strength of forces that favor monopolies of dissemination held in a few hands. Powers that undergo forced industrialization use information control in order to undermine local sentiment, and to encourage national identity. National not transnational perspectives are fostered. We now see that the initial effect of an expanding civilization of science and technology is not the universalization of a world outlook. Hence it has been suggested that heroic measures be taken in the hope of overcoming the parochializing and devisive consequences of information unbalance along national lines. Drastic proposals include a world communication network that furnishes news to all people everywhere, and a world school board that commands attention of all school children for at least an hour a day.

It is to be predicted that wherever popular government is well established and competitive electoral propaganda is important, research information about public motivation will grow more refined. Unless voters and officials are provided with an up to date stream of information about how they are manipulated, popular government will be deprived of the degree of choice that they might otherwise exercise.

In a data rich civilization the most serious constraints on information are likely to be in processing and use. There is danger that modes of perception will be standardized and that motivations will be turned toward unimaginative channels of expression. As mass society grows the leadership will no doubt become aware of the latent threats to established order. Hence it will be tempting to manage education for purposes of rigid indoctrination.

Two broad lines of policy regulation are conceivable in regard to information: limit gathering and storing; limit access. It is most improbable that initiatives toward the former policy line will be accepted. The latter, however, does come within the scope of likely action. For example, public regulation is likely to define who may obtain access to personal data.

It is generally agreed that great damage is done to children if derogatory information is made public about them individually. Attempts are made to avoid reporting the stigma of illegitimacy, of early delinquency, or of initial failure. Adults are also given the benefit of some non-disclosure policies, and no doubt pressure will continue to be exerted to screen health and other records from miscellaneous hands.

However, I suspect that attempts to protect human dignity by the selective enforcement of privacy will almost certainly become obsolete. Privacy won't work. Coming years may witness a drastic new approach for the purpose of protecting our self esteem. Instead of relying on privacy we will shift to strategies of insight. A comprehensive stock of social information will provide facts that reduce the false pride of an individual, family, or community. If the educational system trains the individual to see his enormous potentialities and to discover the many ways of overcoming specific negatives, the resulting character systems will be able to take or leave privacy without grave inconvenience.

Unless the storage banks of society are made accessible to competiting organizations strong enough to utilize and interpret great amounts of data, the monopoly exercised over enlightenment by privileged organizations will throttle the expansion of knowledge. Furthermore, unless provision is made for continuing insight into perceptual limitations and also into restricted motivations, the information potentials of society will remain poorly adapted to the requirements of science and policy.

AREA IV. INFORMATION NEEDS OF SOCIETY

Symposium B: Specific Knowledge Areas

INFORMATION PROBLEMS IN THE FIELD OF ANTHROPOLOGY. Stephen T. Boggs, American Anthropological Association, 1530 P St., N.W., Washington, D.C., 20005, U.S.A.

The study began with producers of information to determine whether a "backlog" of unpublished results of research existed. Questionnaires were sent to all professional members of the Association and 18 per cent replied without follow-up. Interviews were conducted with 80. It was discovered that a very large proportion of the research conducted over the past twenty years is unpublished and consists largely of field notes, which are usually unintelligible to other workers and unevaluated as to context and significance, and data reports, which involve the organization of data into descriptive categories. There was virtually no backlog of unpublished interpretative studies, i.e. those in which selected data and generalizations are related to the working concepts of the discipline. However, there was great (and unanticipated) dissatisfaction among users of information with the dearth of published descriptive data on which interpretative studies are supposed to be based. This was reflected in the strong feeling that "I want to check others' observations against mine and make my own conclusions." What has apparently happened is that prestige, and opportunities to publish, have gone to the interpretative study, while "mere" descriptive reports have been scorned. There is, however, scepticism about the quality, and even the existence, of descriptive manuscripts, as contrasted with unanalyzed data. A follow-up study is needed to pinpoint the nature of these unpublished materials and their rate of publication. As a field anthropology lacks great quantities of repetitive observations, such as time series. It needs conceptual flexibility and open-endedness in any information system and probably a two-stage retrieval operation: one consisting of an index of all possible topics and the other something like literary texts, which may never be put into a machine.

INFORMATION EXCHANGE PROBLEMS IN PSYCHOLOGY. John G. Darley, Chairman, Department of Psychology, University of Minnesota, Minneapolis, Minnesota.

A principal difficulty in psychology stems from the confounding and confusing of archival functions and information exchange functions. Publication traditions and structure grow from our archival history. This produces a tight interaction net comprising the linked variables of: budgeted pages/dollars per journal; number of manuscript submissions; rejection rates; constraints on article length; form of editorial compensation or term of office; and publication lag. When we consider the behavior of producers and users of knowledge, identifying small groups of producers and several user groups such as students, teachers, practitioners, other researchers in the "invisible colleges," secondary source producers (abstract services, annual review authors, or textbook authors), it becomes apparent that the direct readership of a single archival item is small, for the total effort involved in its production. There is clear evidence that a parallel information exchange structure has emerged, however, that compensates for shortcomings in the archival structure but increases the redundancy of the total system. This exchange structure includes at least the following elements: preprints; government technical reports; velocity of travel and consulting relations; colloquia; regional and national meetings; and "in-groups" in specialized areas. Applications of computer technology that fail to take account of this parallel information exchange system may be both ineffective and excessively expensive. Innovations require reconsideration of the editorial "gatekeeper" function and creation of procedures to enhance exchange, such as: convention proceedings publications; rapid abstracting plus item purchase systems; appropriate identification of classes of users. Grouping of disciplines to which these findings may be relevant awaits further study; it is probable that discipline differences exist in epistemological practices, information exchange needs, and training programs at the undergraduate and graduate levels.

INFORMATION NEEDS IN EDUCATION. John K. Hemphill, Educational Testing Service, Princeton, N.J., U.S.A.

The "explosion of knowledge," of which we are now becoming only too well aware, finds education squarely in its path. Whether the explosion can be contained in such a manner as to permit it to perform useful work depends upon our ability to cope with some newly emerging problems. How can large masses of new knowledge be organized in a manner that will permit the individual to orient himself effectively to its existence? It seems no longer possible to expect any man to be aware of other than a very small and highly specialized segment of existing knowledge. Today's best insights are rapidly being replaced by superior insights. Sheer quantity and rapidity of change in man's knowledge creates unusual problems for curricula at all levels of education. Does information handling technology contain within it the seeds of a vastly more efficient method for transmitting the knowledge of one generation to the next? This question goes beyond passive "storage and retrieval" of information to active utilization of what we know. We see some indication of a potentially more efficient educational technology in present concerns with "programmed instruction" and "computer-assisted instruction." Our educational enterprise is one of the most loosely controlled in existence. Its efficiency is unknown and its effectiveness largely a matter of faith. The educational enterprise is in dire need of feedback information of all kinds if it is to meet future responsibilities. How well educated are the people of our country as they leave our schools? A national educational assessment program is seriously being planned. How well prepared are our teachers, our principals? How can the information needed for effectively planning federal aid to education be provided? Are there differences in the quality of education received by differing segments of our society? Are we permitting the talents of capable but unidentified individuals to be wasted? All these questions and many more are being asked today. Each may have an answer well within the grasp of efficient information handling technology.

SKETCH OF A PROPOSED SEMI-AUTOMATIC, HIERARCHICAL, OPEN-ENDED STORAGE AND RETRIEVAL SYSTEM FOR STATUTE-ORIENTED LEGAL LITERATURE. Layman E. Allen, Yale Law School, New Haven, Conn., U.S.A.

In that aspect of the social process called law the reasonable expectations of individual human beings are of central significance. Since such expectations are dependent upon the fund of information available, methods of storing and retrieving legal literature, as well as the modes of expressing the ideas contained therein, profoundly effect aspects of human life touched by law. It may be that the normative character of statements in legal literature makes a "language normalization" approach particularly appropriate for the storage and retrieval of such statements. To date such possibilities of relating language normalization to the handling of legal literature are unexploited -- in fact, they are virtually completely unexplored. One approach for normalizing legal literature to facilitate its organization for efficient storage and retrieval is considered here. This approach is in a preliminary stage of development; thus, it will be sketched in broad outline only. Normative statements can be transformed into a normal form of the following character: If certain specified conditions are fulfilled, then certain legal consequences follow. The normalized version of a normative statement is, thus, an implication whose antecedent is a set of conditions that are conjunctively and/or disjunctively joined and whose consequent is a set of legal consequences that are conjunctively joined. Normalized versions of the central ideas expressed in a document can be related to a Normalized Sentence-Index-Matrix (N-SIM), which will serve as the means for relating that document to all other documents that deal with that same subject matter. The N-SIM approach to storage and retrieval of normative statements is characterized by the following features: (1) the indexing categories are complete sentences, rather than words or phrases; (2) the indexing categories are both hierarchically-ordered (in a way that permits both human and machine processing) and open-ended (so that there is no need to guess now what categories will be relevant in 1984; rather the evolving literature determines new categories to be added); (3) at present it is aimed at statute-oriented legal literature only.

THE INFORMATION NEEDS OF POLITICAL SCIENCE.

Karl W. Deutsch, Yale University, New Haven, Connecticut, U.S.A.

Political Scientists need three main kinds of information: 1. current data on office holders, legislators, recent votes and other events where speed of information is essential; 2. background data from economics, psychology, sociology and other disciplines, for current or past events, where comprehensiveness and accuracy are essential; and 3. time series and historical data for past crises, decisions, recurrent configurations and long-term trends. Beyond the raw data, Political Scientists need information on their margins of error of these data, and of the permissible errors in using these data in the context of particular theories or efforts at prediction. In the unfolding dialog between theory and data--such as at the Yale Political Data Program--the deeper statistical and mathematical analysis of data is becoming as important, or more important as is the mere gathering of primary data. Data archives are thus changing into data laboratories, and computing, analytic and retrieval facilities should be designed accordingly.

INFORMATION NEEDS OF SOCIETY WITH RESPECT TO THE AREA OF HEALTH.

Marcus Rosenblum, Department of Health, Education, and Welfare, 330 Independence Avenue, S. W., Washington, D. C., U. S. A.

With the caveat that professional needs [illegible] necessarily social needs, significant social needs may be characterized as follows: acquisition, organization, [illegible] management of data, methods of marketing information [illegible] market demands. The first of these three classes [illegible] needs is identified by proposals [illegible] The second is a response to both [illegible] market demands for available information [illegible] articulated by professional health workers [illegible] market is discerned in morbidity data, [illegible] and type of common misinformation, data [illegible] from [illegible] and conspicuous behavior patterns. To satisfy such needs and to respond to persistent and characteristic demands of information users, society wants an extensive system of information outlets linked to specialized centers and universal information depots. The rudiments of such a system have been constructed almost inadvertently; with plans and money, all demands might be satisfied. Unlike most parcels of scientific information, health information must be widely distributed; a strong economy depends on it. As the technological needs of a health information system have been well anticipated, the main challenges are directed to the social and intellectual resources. These challenges have been articulated in a spate of recent reports on education, medical schooling, drug information, and governmental communications.

INFORMATION NEEDS IN THE FIELD OF ECONOMICS

Edgar S. Dunn, Jr., Resources for the Future, Inc. 1755 Massachusetts Avenue, N.W., Washington, D.C., U.S.A.

In recent years research in economics is moving in two directions that require new information resources and place new and unusual demands upon existing information sources. First, the most pressing social problems of the day have to do with structural phenomena at quite disaggregated levels of detail. This contrasts markedly with the earlier preoccupation of the profession with economic stability as measured by aggregate indicators. Second, economists are becoming increasingly dissatisfied with static state equilibrium models and are seeking new ways to structure their concepts of reality. Stochastic or probability models and systems analysis mark this trend. There is also a trend toward considering a broader set of significant variables extending beyond earlier concepts of the boundaries of the discipline. These trends have been accelerated and reinforced by the advent of the computer which has made available a new order of economy and technical research capacity. These developments have underscored sharply the inadequacy of the systems that generate and make available existing social science and economics information. There are anomalies on both the demand and supply side of the information process. For a number of reasons these problems are qualitatively different from those characteristic of the physical sciences

AREA V. PRINCIPLES OF DOCUMENTATION AND SYSTEMS DESIGN

Symposium A

SOLUTION TO INFORMATION PROBLEMS OR OBSOLESCENCE? Dr. Eugene B. Konecci. National Aeronautics and Space Council, Executive Office of the President, Washington, D. C., U. S. A.

Accurate and timely information is the lifeblood of any organized system, be it government, management of a corporation, scientific knowledge, or intelligence. From the earliest of times man has been dependent on a means of communication between himself and his immediate family, his tribesmen, and even communicating with domesticated animals. Since communication is not a one-way process, even the early cave man had to learn to distinguish not only sounds like a growl, or later, words which form the first vocabulary, but in fact had to use his eyes in the same way that we use ours for determining facial expressions, friendly or hostile acts by other cave men or animals. A greal deal of hereditary instinct such as the need for food, survival, played a large part in the motivation and desire for communication with other primitive people. The more cultured and advanced societies of our past history, such as the Chinese, Egyptians, Greeks, Incas, Romans and so forth, established a communication and information system to a high degree. Mathematics was founded and based on symbols. People trained in the field were able to communicate meaningful information. These facts and knowledge have been winnowed and sifted through generation upon generation for several thousand years of civilized existence of man.

Today we come to a steppingstone in history in the area of information processing, storage, retrieval, and so forth. It becomes quite apparent that with our scientific and technological advances we have become a world that could be buried in megatons of paper. The written word has become an important symbol of prestige and earning power. It is the socially accepted manner of communicating information on a wide scale. The scientist, the systems manager, the executive, and many others are drowning in this mass of ink and paper. Creativity, free thought and progress are being held back by our obsolete methods. The technological solution to this problem is available. However, people in various walks of life are very reluctant to change the accepted standards of newspaper, magazines, journals, and textbooks as being their forms of documentation and communication. This paper will discuss the problem, and alternatives.

Computers, data processing equipment, systems analysis techniques and sophisticated human engineered display and control equipment can help us solve the mechanical aspects of the information problem. But to maintain a high degree of scientific, technological and social progress we will have to improve our education in and out of formal institutions of learning in order to upgrade our people. This means we will have to very well understand how the human mind works, in other words foster the new field of Biocybernetics. Information and documentation is not an end unto itself. It is a requirement in our evolutionary progress. Goods are plentiful and cheap - where labor, skilled labor is expensive and scarce. We must have information systems to continuously and dynamically upgrade man's skills to their fullest capacities and not permit human obsolescence.

PRINCIPLES OF SUBSTANTIVE ANALYSIS OF INFORMATION. Vladimir Slamecka
School of Information Science, Georgia Institute of Technology, Atlanta, Georgia, U.S.A.

In the context of this paper, substantive analysis of information is understood to be an intellectual or intellect-imitating activity serving the input function in information storage and retrieval systems. A model of this activity is developed permitting to differentiate between the process of information reduction, as in abstracting, and information categorization, as in indexing. Emphasizing the state of the art of operating systems rather than of hypothetical or experimental ones, the principles of substantive analysis of information presented are of the nature of essential, underlying constituents, and of rules of action. The principle of relativity of information is introduced as a fundamental constraint in the process of substantive analysis of information. Based on this principle, generalized rules of action, employing predominantly theorems and concepts of mathematical logic and information theory, are postulated for the activity of information representation and categorization, and for information structuring (the formalization of intellectual relations between information classes). Some implications of these principles on the design of information retrieval systems are discussed.

THE ROLE OF PAPER TAPE AND OPTICAL SCANNING COMPUTER INPUT IN TEXTUAL DATA PROCESSING. Raymond P. Wishner. Documentation Incorporated, Bethesda, Maryland, U.S.A.

For many years to come the search system of a mechanized library will interrogate and manipulate for indexes surrogate records rather than entire documents word by word. Many aspects of the creation and maintenance of a document surrogate may be likened to creating and maintaining a stock record in inventory systems found throughout mass production industry. Unlike the maintenance of an inventory status system is the necessity that many entries of the document surrogate record must be nearly free form text. Examples of this are the document title and abstract. Whereas in a standard file maintenance system the form rigidity of punched cards is welcome discipline to the systems designer, the rigidity of the punched cards is an uneconomic constraint for "inputting" free text. This paper will not fight the ten year running battle between paper tape and punched cards as the basic input medium for a digital computer. Rather it will use the argument of the superiority of paper tape for the inputting of "free text" data into a computer as a bridge to the argument of the potential superiority of a rigid font optical scanning system for the inputting of "free text" data in a mechanized library.

The argument of paper tape versus punched cards as an inputting mechanism has been fought out in terms of rates of key strokes and the character reading rates of computer input devices. In relation to free text inputting, however, these arguments are really secondary to the advantages of textual correction techniques of one medium over the other. In this area, systems designers have often discarded the most important advantage of a paper tape typewriter by using correction procedures well adapted to punched cards but quite senseless for a paper tape oriented system. The two most common paper tape correction procedures are: 1) having an operator blank out a paper tape from the point of error and rekeying the text from this point, 2) a character by character verification of the text by rekeying the data. The object in both methods is to produce nearly perfect input data to the computer in one operational pass.

An alternative and powerful correction technique utilizes the visual copy created in the process of keying the text on a paper tape typewriter. As the data is being keyed, the operator may semantically recognize an error in what he is directly keying or for that matter in a previous paragraph. By giving the operator a base point (i.e., start of last paragraph) and a set of correction codes (i.e., delete, insert, and replace) he can make the correction immediately or delay it to the end of the current paragraph. Regardless of the details of the correction scheme an unambiguous correction can be coded for computer batch processing and a record of the correction made for later visual verification. Further, there is no reason that correction procedures involving the visual copy should stop here. The visual copy can be removed from the typewriter and sent to a proofreader and corrections annotated on the input proof copy. The proof copy can be re-inserted into a tape typewriter and the corrections keyed into an exception paper tape reel. As long as the base record and base point within the base record are clearly coded, an unambiguous correction for computer manipulation again can be created and a record of the correction made for visual verification.

Note must be taken that initial input proofing and correction is often not sufficient in textual data processing. Thus it may be necessary for the computer to print out visual copy for paragraph verification. The same procedure used for initial input proof and correction can be used for computer printed proof copy. . .the advantage is naturally one "well-defined" system for editing.

At this point part of the basic thesis of the paper will be now clearly delineated: The advantages of using the initial proof copy created in keying the paper tape codes as a correction medium are:

1. The ease with which the operator can create corrections for immediate errors and errors created earlier in the keying process.
2. That the proof copy can be removed from the keying machine and later inserted again for the creation of corrections.
3. That the same system used for creating initial textual corrections is nearly the same as the system for correcting computer printed proof copy.

The significant disadvantage to the above described input proofing and correcting scheme is the potential physical separation between the visual proof copy produced by the typewriter and the paper tapes that contain the original copy and later corrections. Systematic handling procedures can negate most of this disadvantage (i.e., batch processing procedures). It is here, however, that the significant advantage of rigid font optical scanning techniques are most apparent.

With an optical scanning system for inputting textual data, the proofing medium and the inputting medium are one and the same thing. Thus a typist using a typewriter equipped with optical scanning font and subject to no more format discipline than with a tape typewriter can create initial input for a computer. This copy can be removed from the typewriter and blue penciled by a proofreader for corrections. The copy medium can then be re-inserted into the typewriter and corrections typed at the bottom of the form. These corrections are coded to refer to some base point in the original input text or in an earlier correction line.

The final thesis of this paper is that rigid optical scanning systems reduce the complexity of the textual data inputting systems by creating an absolute one to one correspondence between the proof copy and the input medium.

Although the price of rigid optical scanners has decreased significantly in the last few years their cost still does not justify their use on small volume library applications. However, as it has been foreshadowed in the argument of this paper, the basic system for using paper tape equipped typewriters and optical font equipment are nearly the same. Because of this close relationship between the two inputting systems, we are experimenting with a data inputting system compatible with both methods.

AREA V. PRINCIPLES OF DOCUMENTATION AND SYSTEMS DESIGN

Symposium B

FILE STRUCTURE AND SEARCH. Terry R. Savage. Datatrol Corporation, Silver Spring, Maryland, U. S. A.

The purpose of this paper is to examine the characteristics of file structure and search techniques in information retrieval systems. The main concentration will be upon the problems and possible solutions available for file structures and searching procedures on general purpose computers. The paper is given in terms of specifications that would be encountered if one were faced with the problem of designing a file and its associated search program for an application in information retrieval.

The first section concerns itself with the actual content of the file. That is, what specific pieces of information must be maintained in the file for each item. Two criteria are employed which jointly provide necessary and sufficient conditions for successful utilization of the file. The first of these is to determine the minimum information required for each item so that it might be matched to any question likely to be posed. Even though the particular pieces of information will vary from system to system, experience over the last few years with very large systems, as well as the folklore of traditional librarianship allows one to generate a sufficient minimum basis for all systems. The second criterion to be applied is that the items in the file must contain sufficient information in order to allow a given searcher or user to determine that a selected item is or is not relevant to his interests.

The second section of the paper concerns the relationship between the physical and logical organization of the file items themselves. The limitations and capabilities of existing general purpose computer systems effect the physical organizations that are possible. Alternative logical organizations are considered in the light of their effect on the economics of file manipulation. All known physical and logical organizations are described and a tentative selection is made of each for the purpose of the subsequent discussion.

The third section of the paper concerns an analysis of the various alternative procedures that may be used to match a given question to a given file item. Standard Boolen, associative, and probablistic matching criteria are described and one new alternative is suggested. It is important to note that the matching criterion is independent of both organization and search strategy.

The fourth section of the paper concerns itself with the alternative search strategies that may be employed, given the selection of the file content, file organization, and matching procedure. It is seen that the selection of these prior alternatives, especially with respect to organization, severely constrains the available choice of search strategy. In the light of this constraint the problem of file organization is reconsidered and alternatives previously rejected are examined in more detail.

The final section of the paper deals with the options available for output of the search system. In particular, systems which include so-called on-line capabilities are discussed in detail. In this section it is argued that satisfactory

selection of output format and content is critically dependent on vague and relatively unknown psychological characteristics of the searcher and the user. Some suggestions are made for increasing our knowledge of users.

Selected References

1. Becker, J. and R. M. Hayes, Information Storage and Retrieval: Tools, Elements, Theories, Wiley, New York, 1963.

2. Bracken, R. A. and H. E. Tillitt, "Information Searching with the 701 Calculator," JACM, v. 4, n. 2, (1956).

3. Libby, R. L., "Mechanization of File Storage and Access," in Proceedings of the Conference on Libraries and Automation, Library of Congress, 1964.

4. Patrick, R. L., and D. V. Black, "Index Files: Their Loading and Organization for Use," in Proceedings of the Conference on Libraries and Automation, Library of Congress, 1964.

5. Schwartz, E. S., "A Dictionary for Minimum Redundancy Encoding, "JACM, v. 10, n. 4, (1963).

6. Seidel, M. R., "Threaded Term Association Files," Datatrol, IR-14, Presented to the NBS-ADI Symposium on Statistical Association Methods for Mechanized Documentation, March, 1964.

7. Swid, R. E., "Linear vs. Inverted File Searching on Serial Access Machines," ADI Proceedings, 1963.

DISSEMINATION OF INFORMATION AND PROBLEMS OF GRAPHIC PRESENTATION. Charles J. Austin. National Library of Medicine, Bethesda, Maryland.

An information system is no better than its final products. The characteristics of good output from a computer system include: simplicity, readability, compactness, completeness, reproducibility, economy, timeliness, and proper sequence. Output products may be characterized by form (printed, other visual, machine-readable), by size of audience (individuals, large groups), and by response time (immediate, rapid, delayed).

Equipment Review. A wide variety of devices are available for producing output from computer systems. Printers are the most common. Electric typewriters, the slowest kind of output printers, operate at speeds of approximately 10 characters per second. Electro-mechanical line printers (which operate at speeds of about 150 to 1500 lines per minute) fall into four categories: type-bar, type-wheel, wire-matrix, and rotating-drum. Chain printers are more flexible, since they permit expansion of the character set at some sacrifice of speed. Electrostatic printers are much faster, operating at speeds greater than 5,000 lines per minute. Photocomposition devices produce film copy of graphic arts quality which can be used directly for plate-making and subsequent publication. Other types of visual output include graph plotters and visible displays.

Some retrieval systems produce machine-readable output for use at other establishments on local equipment. Such output may take the form of punched cards, paper tape, magnetic tape, disk packs, or removable magnetic card packs. Other systems produce remote output by means of data communication links to console units. Finally, the actuator (a digital to analog converter) is an output device which enables a computer to control the operations of machinery and processes. So far, actuators have generally had industrial applications, but they may become important in the information field for linking digital and graphic retrieval systems.

Systems Considerations. Designers of an information retrieval system must have the answers to several basic questions in determining output requirements. Who needs the information? How quickly does he need it, and in what form? Why does he need the information? What does he do with it?

Output requirements directly affect the input to a system. The system designer must ensure that all needed data are entered into the computer. This always includes the substantive data to be printed in reports, and may also include control data, such as typographic symbols for photocomposition and special sorting elements for publications. Output requirements should also be considered in optimizing internal file construction for a system.

Special attention must be paid to designing an adequate character set for a published index with due consideration to problems of special characters and diacritical marks used in various languages. Other important design considerations are the high-lighting of "scan" elements; margin justification; and the use of multiple columns, running heads, and page numbers. The system designer must also weigh the advantages of direct access or index access to a bibliography and give careful thought to the frequency of cumulation.

In developing individualized output products from a retrieval system, flexibility of both format and sequence is desirable. One particularly useful technique is to list citation references in sequence according to a weighted scale based upon the numbers of "hits" of search terms and index terms.

There are special design considerations when providing machine-readable output for others. External users impose limitations on the parent system--i.e., a change to the system has multiple effects. Complete program documentation is essential, and training and technical liaison should be provided. Effective procedures must be developed in advance for maintenance of files released to field users.

Computer manufacturers provide "software," or generalized programming routines, which can be useful in developing output systems. Sort and merge routines linked with report generators can simplify the programming task considerably. Data processing compiler languages are particularly useful in facilitating changes to output records.

This paper presents a detailed review of output equipment for automated retrieval systems, and describes some of the more important elements of system design to be considered.

Bibliography

Chapin, Ned. An Introduction to Automatic Computers. Princeton, N. J., Van Nostrand, 1963, 525p.

Computer Typesetting Conference, London University, 1964: Proceedings. London, Institute of Printing, 1965, 245p.

Sparks, David E., Berul, Lawrence H., and Waite, David P., Output printing for library mechanization. In: Libraries and Automation; Proceedings of the Conference on Libraries and Automation held at Airlie Foundation, Warrenton, Va., May 26-30, 1963, Washington, D. C., U. S. Library of Congress, 1964, p. 155-190.

CONTRIBUTED ABSTRACTS

EDUCACION DE DOCUMENTALISTAS.-
Ermelinda Acerenza, Bibliotecaria de la Facultad de Ciencias Económicas y de Administración. Montevideo, Uruguay.
Muy diversas opiniones sobre formación de documentalistas. ¿Es una profesión con fisonomía propia, o bibliotecario y documentalista tienen bases comunes? Ambos proporcionan elementos de información. Reunen, organizan y conservan documentos, ponen al alcance de los usuarios. Estas operaciones varían, si son libros o documentos derivados de los mismos (catalogación y resúmenes analíticos, estadísticas, informes, estudios, traducciones, extractos). Los programas de las Escuelas de Biblioteconomía incluyen la enseñanza de los puntos señalados, no obstante el egresado carece de preparación adecuada para los fines de un servicio de investigación. En Europa se pretende sentar bases de una profesión, estableciendo cursos para Ayudantes de Documentación y para Documentalistas especializados en archivos o en bibliotecas. Se dice que el bibliotecario especializado cumple funciones de documentalista, tal como éste las realiza en un Centro de Documentación. En Mesa Redonda de la 26a. Conferencia G. de la FID, se dijo: La enseñanza de la documentación puede hacerse en Escuelas de Biblioteconomía por considerarla una especialización dentro de la profesión. Es evidente la necesidad de formar bibliotecarios especializados (en distintas ramas del saber o documentalista)
Conclusiones: Establecer en Escuelas de Biblioteconomía cursos especializados de postgrado con estas asignaturas: Selección científica y técnica; bibliografía especializada; análisis, resúmenes y extractos; organización de materiales especiales y de la información; clasificaciones bibliográficas; técnicas de O&M; normalización; reprografía; sistemas de análisis conceptual para la localización automática; servicios de investigación; instituciones para el desarrollo de la información. Que las Escuelas de Biblioteconomía concedan al egresado del curso de postgrado, el título o certificado de Bibliotécnico Especializado (en alguna rama del saber o documentalista).-

NEW DEVELOPMENTS IN DRUG INFORMATION. **Boris B. Amzlower, Pharmaco-Medical Documentation, Chatham, New Jersey 07928 USA**

Drug information needs and flow patterns have changed radically in the two decades since the then revolutionary wartime crash program in pharmaceutical research and development brought cheap penicillin as one of the most spectacular and tangible benefits. The worldwide acceleration of pharmaceutical research, with the continuously rising production and consumption of drugs has made it increasingly difficult to service effectively drug information needs using classical tools and techniques; this difficulty has been compounded by sudden massive developments of totally new classes of drug use (e.g. oral ovulation inhibitors) or of unexpected types of drug toxicity (e.g. interference with specific phases of embryonic development). A recent study based primarily on UNLISTED DRUGS, a cooperative international drug-identification service established by the Pharmaceutical Section of SLA in 1949, revealed the presence of several key trends; a number of these can be extrapolated to show probable future drug information needs and some indications of most efficient methods to satisfy them. For individuals and organizations concerned with either drug effectiveness and safety or drug development and marketing, an improved definition of common denominators in the field of drug information should bring abundant returns for the efforts expended.

A SYSTEMS APPROACH TO DATA BANKS INCORPORATING BOTH MANAGERIAL AND TECHNICAL DATA. K. B. Andrews, A. K. Dunlap, and M. N. Sloane, TRW/Space Technology Laboratories, One Space Park, Redondo Beach, California, USA.

A systems approach to the interface problems of managerial and technical data in future international data banks is presented. The impact of the scientific and technical data banks in purely managerial areas such as law, economics, and marketing is clearly evident today. The timely flow of information between the two fields is of the utmost importance. At the same time private proprietary rights and governmental interests must be preserved. A model of an international oceanographic data network containing managerial information is discussed in some detail. Systems-type flow diagrams are presented showing the interfacing of managerial and technical data as well as charts to show the hierarchy of levels of information in the network at international, national and regional levels. The relationship between the oceanographic data bank and other data banks which must feed data into the system such as weather information, food and agriculture, etc., are pointed out. A built-in statistical screening procedure for keeping data in the system current and relegating other data to historical files as various projects phase out is discussed. Oceanographic research is on the verge of opening up vast new areas in mining, ocean-agriculture, law, marketing, etc., all of which add to the complexity of management problems.

A PHARMACEUTICAL DATA HANDLING SYSTEM FOR SCIENTIFIC AND MANAGEMENT REPORTING FUNCTIONS. Frances H. Arendell, Carol R. Lengler, Cynthia L. Goebel, Dr. Lee N. Starker, Warner Lambert Research Institute, Morris Plains, N. J., U.S.A.

Large numbers of synthetic organic compounds must be screened by preliminary categorizing tests, and a few subjected to intensive animal study to produce one clinical candidate. This process results in large volumes of data, much of it negative, but most of it in consistent formats. Unfortunately, this data never dies, but must be re-searched for compounds having certain patterns of activity newly found to be a meaningful lead for useful clinical properties. The biological data can fall fairly readily into three types: 1. Animal dose-response data from fixed testing procedures, amenable to standardized reporting formats. 2. Scores of subjective observations for increased or decreased behavior, or blocked or potentiated challenge drug responses, for which score sheets can be used. 3. Evaluation data from non-routine testing procedures devised or modified to bring out the particular properties of a drug in advanced development, reported in technical reports. Basic tab card formats were designed to handle these three types of data in an integrated system using an IBM 082 sorter and 870 document writer to start. All data may be listed by compound, or each test reviewed for the most active compounds. Simple fixed programs may be used for fairly effective tabular displays of the data, with or without Markush variants of chemical types. Quick search cards for screening status reviews or pattern searches can be generated from the master cards, using one column per test. A series of management reports were devised covering active compounds (permitting a pre-edit of the report by the scientific staff), quantities of tests run per month, project surveys, etc. We have proven the feasibility of our concept of a unified reporting system from the single input to the data store on a card handling basis over the past two years. The present size of the operation and the data store warrants conversion to magnetic tape.

SYNTOL–A NEW SYSTEM FOR THE ORGANIZATION OF INFORMATION. Susan Artandi, Graduate School of Library Service, Rutgers, the State University, New Brunswick, N.J., U.S.A.

SYNTOL is examined in its capacity as a distinct system for information organization primarily from the point of view of the type of indexing involved and from the point of view of the device, built into the system which condition the type of searching that is possible. The objective is to see how SYNTOL differs from or is similar to other systems and to determine what features are unique to SYNTOL. As a post-coordinate system at the word level SYNTOL corresponds to a coordinate indexing system which does not use roles and links. At the syntagmatic level the combined effect of roles and links is achieved through the use of highly formalized relations between terms and through the orientation of these relations. Individual syntagmas can be also regarded as pre-coordinate indexing terms in which two descriptors are linked together with no indication of relationship between them or in which some general relation is indicated between the terms which make up the syntagma.

[illegible] [illegible] [illegible], American Institute of Physics, [illegible] Street, New York, New York [illegible].

A certain [illegible] of overlap (or interface) [illegible] between [illegible] publications in a special [illegible] to cover its field of interest. The dimensions of [illegible] overlap for [illegible] journals in three abstracting [illegible] is displayed [illegible] on an article-by-article coverage [illegible] with particular reference to the coverage of *Physics [illegible]*. This quantitative study is based on data collected [illegible] AIP-IRP [illegible] study of *Physics Abstracts*, Institute for Scientific Information's report on article-by-article coverage of selected abstracting services, and Aslib and AIP-IIP comparisons of abstracting publications and specialized bibliographies. The journals studied are grouped into four categories:

(a) Major physics publications
(b) Multi-disciplinary publications which contribute significantly to the physics literature
(c) Multi-disciplinary publications which make a small contribution to the physics literature
(d) Major journals from other disciplines which publish a small number of physic articles.

In addition to showing the article-by-article overlap, this paper also gives information on the number of articles unique to any one of the abstracting services covered in this study, together with details of the articles that were not abstracted.

ANALYSIS OF INFORMATION NEEDS OF INDIVIDUAL SCIENTISTS WORKING IN THE RESEARCH ORGANIZATIONS OF THE RUMANIAN ACADEMY. Pia Atanasiu and Georgeta Lăzărescu, Scientific Documentation Centre of the Rumanian Academy, Gutenberg St., 3 bis, Bucharest, Rumania.

An analysis of methods utilized in investigating the users' needs and a review of world-wide results obtained up-to-date are made. Using the inquiry method by means of a questionnaire distributed to 500 rumanian academic research scientists, the authors establish the preferences, uses and individual information gathering. A critical comparative study of similar data obtained abroad is done. The study intends to draw up conclusions concerning both the efficiency of the investigation methods for establishing the users' needs and some local specific aspects, the latter determining differences in information gathering behaviour.

CURRENT INFORMATION PUBLICATIONS. CONSIDERATIONS ON THE EXPERIENCE OF THE SCIENTIFIC DOCUMENTATION CENTRE OF THE RUMANIAN ACADEMY. Dr. Aurel Avramescu, member of the Rumanian Academy. Scientific Documentation Centre of the Rumanian Academy, Gutenberg St., 3 bis, Bucharest, Rumania.

The usual types of current information publications and their efficacy are compared. The paper points out the necessity of providing bibliographic references both with UDC numbers and descriptors, for unspecialized journals. In author's opinion, although classification and indexing require more time and delay publication, they yield better informational value and wider possibilities for retrieval. Solutions adopted by the Rumanian Documentation Centre and measures taken for improving initial type of its current information publications are discussed. An increase of the descriptors number assigned for each paper and the utilization of automatic indexing methods are contemplated.

[illegible] preprints [illegible] reviewing process is [illegible] and far from comprehensive, one seeking to find a literature of experience in this effort is ill-served by the prevailing pattern of secrecy. Poor reviews. A system of screening the published and report literature in the area of applied thermal processes is described which leads to a select listing of citations [illegible]. These, plus the cited references of the articles published papers, are key-punched on cards, and a listing of the cards is provided as copy to the printer, in lieu of galleys, for a General Bibliography, serially numbered. The punched cards serve as input data for preparation of a KWIC Index of the entire bibliography. Citation Indexing for the original published papers is made possible through an editorial identification of cited references already coded in the general bibliography previously published. The punched cards are available to be copied for external documentation research.

*Research Consultant, The Marquardt Corporation, Van Nuys, California.

A CLEARINGHOUSE FOR SCIENTIFIC AND TECHNICAL MEETINGS. ORGANIZATIONAL AND OPERATIONAL PROBLEMS. Harry Baum, Technical Meetings Information Service, 22 Imperial Drive, New Hartford, New York, U. S. A.

There has been much interest expressed in the formation of a clearinghouse for scientific and technical meetings. In spite of this interest, no clearinghouse has been formed. Formation will not occur until the thrust toward a solution of the problems involved is great enough to overcome them. Whether or not there is a need for the existence of such a clearinghouse has not been proven, but in this paper I shall assume that the need exists and go on to discuss the problems. These include: (1) definition of area and level of coverage of the subject matter of any given meeting (2) definition of the geographic area from which a meeting will draw its attendance (3) defining the area of interest of the clearinghouse in the face of the difficulties imposed by interdisciplinary meetings and by project-oriented meetings (as distinguished from discipline-oriented) (4) attaining comprehensive coverage in spite of the difficulty in obtaining inputs from the organizers of government classified and other closed meetings, as well as from the organizers of ad hoc meetings. (5) the need to ensure that the organizers of meetings will make use of the clearinghouse and (6) the need to take into consideration meetings outside the United States as well as those inside. Solutions to these problems will not be found by a simple approach, but will be obtainable only as the result of careful study of the structure and dynamics of the international and national scientific and technical meetings network.

AN INTEGRATED EXPERIENTIAL TRAINING MODEL IN DOCUMENTATION. C. D. Batty, College of Librarianship, Aberystwyth, Wales.

To facilitate training in documentation the College of Librarianship, Wales is establishing an integrated set of manual and mechanized indexing systems all handling the same body of prepared information from a limited field. This constitutes a model to be used for training through experience, in three stages: demonstration by the lecturer; practice in use by students (searching); and practice in development by students (indexing). The model's scope will also include prior and subsequent stages in documentation activity. The material used as a basis for the model is the literature on information retrieval, because of the familiarity of the concepts and terminology, the limited size of the field and therefore also of the index languages, and the existence of Library Science Abstracts. Systems already in existence include a conventional classified catalogue using the CRG scheme for library science, a rotated classified file in visible index form, a Uniterm system, a manual co-ordinate system using random coding on edge-notched cards and a mechanised co-ordinate system using positional coding on body-punched cards. Additional or substitute systems proposed for a computer now under consideration include a version of the classified catalogue, a serial file and an inverted file based on the co-ordinate systems above, a KWIC index and a "lattice" index. Supplementary to these systems, which are dynamic in the sense that they will develop through use, is a static set of variant systems for demonstration only, e.g. Peek-a-boo as a variant on Uniterm. The model is "primed" with 500 documents coded in each system for initial demonstration and use; indexing practice by the students will extend and develop the model. It is essential for all systems to be kept at the same stage of development for purposes of demonstration and comparison, though not all systems will be developed by the students. The model also offers experiential training in the preparation of documents (e.g. abstracting) and the dissemination of information. It will also offer some training in SDI systems, and on later stages of its development may act itself as material for research.

ACCELERATING INFORMATION TRANSFER IN SCIENCE AND TECHNOLOGY, Charles L. Bernier, The National Library of Medicine, 8600 Rockville Pike, Bethesda, Maryland, U. S. A.

This is an analysis of: what can be done with existing technology; where resources are required; and what new research & development (R&D) in information science is needed to speed application of results of R&D. Between information generation and use there are many blockages and delays. Resistance to novelty--the heart of successful R&D--is a block. Another block is indicated by delay between reading and application. Delays exist between results and their communication; between writing and publication; and between publication and reading, extracting, abstracting, indexing, reviewing, teaching, and learning. Delays in printing of R&D results invite comparison with newspaper accomplishments. Delays in reading relate to competing diversions of a complex civilization including the increased complexity of R&D itself. New techniques for selecting, condensing, and presenting information (including data) are needed to make salient facts interesting. The potential user needs more attractive leading into reading. New selectors for documents and data are required as well as more extensive use of existing selectors. Measurement of reading and scanning times for different formats are urged. Better abstracts are considered to enable the scanner to select more accurately and rapidly. Inadequate support of information services and systems causes delays. Support includes decision to act, selling necessary ideas to resource sources, and resources (dollars, position allocations, and personnel) to carry out the decision. Backlogs of papers need resources for publication. New periodicals, evaluative reviews, journals of indexed abstracts, and data compilations are needed. Techniques for producing these are discussed. Full application of existing techniques through improved resources can contribute to promptness of communication.

A DISCUSSION OF SOME INDEXING PROBLEMS. Masse Bloomfield and Emil Schafer, Hughes Aircraft Company, Culver City, California U. S. A.

A brief study was made of the indexing approach of several different indexing-abstracting publications. In the sample taken, five different articles were selected. Each article is indexed in Chemical Abstracts, Nuclear Science Abstracts, Engineering Index, ASTIA Technical Abstract Bulletin, Physics Abstracts, and in a KWIC (Keyword-in-context) system, manually prepared. The derived results are discussed and graphically displayed. Widely different patterns of indexing were observed, indicating either a lack of effective indexing rules, or that each indexing journal not only serves a different public, but that each public has little in common with the other. One of the tables developed indicates in detail the syndetic apparatus associated with each set of indexing terms. From this, the individual patterns established by each journal may be ascertained. It is concluded that different systems will not only arrive at different basic vocabularies, but that these vocabularies will express different meanings in each system. The effect of the indexing system on the vocabulary chosen to describe the information content should be investigated further.

DEVELOPMENT OF AN ACADEMIC COURSE ON THE ENVIRONMENT OF TECHNICAL DOCUMENTATION SYSTEMS. Leo M. Bohnert, The American University, Washington, D. C., U. S. A.

The course is based on the assumption that technical documentation systems are best understood as a part of the total communication process among technically trained personnel. Documentation is only one of the regular means of communication in organized research and development work. If the function of documentation systems is to unite members of a specific group of technical users with appropriate documents at suitable times and places, then choices and evaluations of documentary services and techniques can best be made in the context of the particular work environment. Study of the environment, both of the technical users to be served and of the documents appropriate and available, will mainly determine what documentary services are practicable, and what few documentary techniques are applicable. An American University graduate course, #55.650, "The Types and Uses of Technical Information", was first given from this viewpoint in the Spring Session, 1965. Three main kinds of users of technical documents were distinguished; management, operating, and research and development personnel. The considerable documentation experience of students in the nation's capital benefitted class discussions and resulted in individual student projects evaluating the effects of a particular work environment on certain documentary activities. One of the main difficulties encountered was in securing adequate accounts, verbal or written, of the various types of work situations that condition existing documentation systems.

CUE INDEXING SYSTEM (CIS). Eric H. Boehm, American Bibliographical Center, 800 East Micheltorena Street, Santa Barbara, California, U. S. A

The Cue Indexing System (CIS) has a dual function: 1) the establishment of a new knowledge classification system which is entirely alphabetical and has mnemonic capabilities greater than that of the Dewey System, the Universal Decimal Classification or the Library of Congress System, and 2) creation of a new type of index which applies computer capabilities to print specific subject facets--the cues--with each index entry. The main cues are 3-letter abbreviations, such as FOR for "foreign relations," or acronyms such as CMP for "communications, mass media or publishing." The index carries the cues with each number entry. It is conventional in other respects: the headings consist of a complete text and they have subheadings and third-order headings when required. A computer is used for permutation of the cues, for printout and later cumulation, information retrieval, and SDI (selective dissemination of information). The first use of CIS will be for the Index Number of Volume I of AMERICA: HISTORY AND LIFE (1965). Some specimen lines:

Abstract	Cues	Years
LAW		
501	CMP:Press SLA REL	1710
218	ECO:Ind.	1932-45
109	ECO:Ind.(steel) LAB	1933
17	SLA REL POL	1860

CIS has general applicability to the special needs of the social sciences and humanities, just as the KWIC index has had general applicability to the current awareness needs of scientific research communities. CIS has these characteristics or advantages: 1) the cues are selected by indexers on the basis of an authority list, 2) the printed text is much less than KWIC, 3) CIS can be scanned more quickly than KWIC, and 4) CIS is not so bulky as to make multi-volume cumulations prohibitively expensive.

DEVELOPING AN INFORMATION SYSTEM FOR MENTAL HEALTH: A PROGRESS REPORT. Lorraine Bouthilet and Julius Frome. National Institute of Mental Health, Bethesda, Maryland.

The National Clearinghouse for Mental Health Information, (NCMHI), an information evaluation center, is developing a mechanized information storage and retrieval system as part of its total system. At present, an interim system has been developed for the IBM 1401 8K, 4 tape drive computer. Index-code sheets as well as thesauri have been prepared by techniques which call for a combination of extracting concepts from the documents themselves as well as obtaining the assistance of experts in the field. Terminology has been developed in such areas as psychopharmacology, crime and delinquency, mental retardation, drug abuse, and occupational mental health. The indexing system is a coordinate indexing system with provision made for the elimination of false drops inherent in such a system. The retrieval system is based on a combination of inverted and serial files, using the best features of each. The computer system is modular in that it can search both specific areas as well as a complete master file. Over a thousand computer searches have been made on a file of over 25,000 documents, which is growing at a rate of 2000 documents a month. The computer search print out contains both the citation of the document as well as a complete abstract. The computer has generated reports in the field of mental health as well as reports concerning a comprehensive survey of the field of mental health information. Upon analysis of ten months of operation of our interim system, the use of natural language, automatic encoding and formatting, the use of a random access computer, suggests themselves as improvements.

COMPATIBLE FORMAT FOR MACHINE READABLE RECORD OF BIBLIOGRAPHIC DATA. Ritvars Bregzis, University of Toronto Library, Toronto 5, Ontario, Canada.

The development of fast access mass storage devices, associative techniques of handling non-numerical data and flexible input and output devices is creating considerable activity in machine applications to information control. One aspect of automated information control - control of bibliographic information - can already be implemented and can expect rapid further development. Several systems of automated bibliographic control are presently in operation. The practical problem of compatibility between these systems has arisen since such compatibility would facilitate quick and accurate exchange of bibliographic information between libraries and documentation centers. The key factor for such compatibility is a uniform format for bibliographic data to be communicated in machine readable form which could satisfy the common information requirements of the participants. A bibliographic data format based on structurally arranged and individually addressable categories and elements of bibliographic data units (titles) has been developed and implemented for the Ontario New Universities Library Project (ONULP) at the University of Toronto Library. An outline of this structural data format is presented and a revised and more detailed version is described and suggested as a basis for universal bibliographic communication in machine readable form.

ANALYTICAL TECHNIQUES FOR THE REVISION OF THESAURI Dr. D.J. Campbell, Pressed Steel Co.Ltd., Oxford, England.

Present methods of revising thesauri are varied and unsatisfactory. They do not indicate an ideal network of references between keywords, or the existence of synonyms. The author's methods have been applied to an index to 3000 documents, using about 1600 keywords, on car-(automobile-) body manufacture and related fields. This was a trial, since no adequate thesaurus was available, by abstracting terms as far as possible from structured terminology, not from the documents. A hand-sortable card record was written for each keyword, giving all existing references to and from it, and any scope-notes. Every keyword has been analysed a) by facet analysis into things (main classificatory items), materials, persons-organisations, operations (main body), processes-properties-phenomena, spatial-geographic, time-groupings, relationships, and residual concepts; b) by things, materials, etc. implied (as driving implies a vehicle, or solder implies soldering); c) by form of energy (mechanical, chemical, etc.) involved and d) by main subject using our own special classification. Punchings are made for keywords falling into any of these groups to be isolated, and analysed into suitable groups, revealing synonyms or the need for references, so that these can be dealt with to make a more efficient and reliable indexing system. Records of time taken allow costing of the whole operation. Records of the analyses facilitate indexing, retrieval and insertion of new terms. Plans are in hand for tests by Cranfield methods of the index before and after revision, but results will appear in another paper.

THESAURUS CONSTRUCTION - HISTORICAL BACKGROUND AND USE CONSIDERATIONS. **Everett H. Brenner, American Petroleum Institute, New York, New York, U.S.A. and Theodore C. Hines, Columbia University, New York, New York, U.S.A.**

The structure of recent thesauri is analyzed by principles evolved for the American Petroleum Institute thesaurus. Background from cataloging heading lists, their syndetic apparatus, and classification schedules is related to modern thesauri, and use distinction is made among subject cataloging, bibliographic unit indexing, and subject indexing with subject headings and with term headings. Some conclusions reached are as follows: the ASTIA, Bureau of Ships, and Bureau of Reclamation thesauri are essentially subject heading lists; the ASTIA and Bureau of Reclamation thesauri contain elements of classification; the Engineers Joint Council and American Petroleum Institute thesauri are more obviously designed for coordination of descriptors; the Engineers Joint Council thesaurus is structured generically in a simply faceted manner with long unstructured lists of related descriptors; the American Petroleum Institute thesaurus purposefully has a faceted system approach which allows automatic posting of generic relationships.

A COMPUTER-GENERATED INDEX PUBLISHING SYSTEM. Brice Carter, Barbara Shaffer and Diana DeWitt, American Society for Metals, Metals Park, Ohio, U. S. A. 44073

The computer-generated published index system, a joint design and developmental effort of the American Society for Metals (ASM), Engineering Index (EI) and the International Business Machine Corporation (IBM) is a generalized and multipurpose program with a wide range of applications and subjects. Designed as a component of the IBM 1401 Combined File Search System (CFSS), it will be included as a module of the System/360 version of the 1401 Information Storage and Retrieval System. The publishing system program arranges all the entries--subject headings, titles, cross-references, authors--in alphabetical sequence, automatically inserts cross-references and prints the index. Abstracts are also handled in this system, and an entire volume of indexes and abstracts, such as ASM's Review of Metal Literature or EI's Plastics or Electronics sections, can be produced as a product of the computer operation. There are five programs comprising the publishing system. The first one extracts desired units from the CFSS Master File, the second does initial formatting, the third generates cross-references from the thesaurus, the fourth sets up page and line format, and the final program produces copy ready for photoreduction. Both the basic format of the CFSS package and the design of the ASM-produced Publishing System allow for great flexibility of index type and content. In addition to the present subject heading and author indexes, many variations are envisioned, including bibliographic citation indexes, listing of notation on content rather than titles and classification by journal. The development, function, utilization, flexibility and description of the publishing system programs are discussed. Detailed program descriptions and system flowcharts are included.

ORGANIZATION OF A NATIONAL INFORMATION SYSTEM BASED ON SUBJECT-ORIENTED DOCUMENTATION. Inge Dahlberg, Deutsche Gesellschaft fuer Dokumentation, Schubertstr.1, Frankfurt, Germany

Awareness of a trend towards a national information system in the United States occasioned organization of the recent 2nd Annual National Colloquium on Information Retrieval (Philadelphia, Pa.). Various proposals for nation-wide systems were offered, most of them based on a centrlized input, the general aim being to provide every scientist in the nation with console-computer access to a centralized or decentralized information retrieval system. However, before any effective information system can be put to work, the necessary documentation of the input has to be organized in advance. A plan for a comprehensive national documentation system is presented herewith. With the help of the National Referral Center, existing information resources could be classified "horizontally" according to subject fields and "vertically" according to depth of analysis, which latter would comprise three stages: special libraries, documentation centers (abstracting and indexing services) and data centers. From this classification a central authority should arrange for subject coverage and designation of types of resources. Among the advantages of such a systematization are: 1) general awareness of institutions responsible for provision of total subject coverage in one particular field; 2) possibility of cooperation with other countries' similarly-oriented centers; 3) availability of all of the information products of the three stages (various indexes, tables, etc.) in specially assigned centers, preferably in the reference departments of universal libraries, rather than in a single, national center, thus providing nation-wide accessability of manual (or eventually electronic) reference tools to universal or specialized collections. Scientists will then have fast and easy access to the total literature in their fields with the new foreign literature adequately analyzed, and they can be tied electronically into such a reference system as soon as the necessary equipment is available.

OPTIMIZING RETRIEVAL RESULTS WITH MAN MACHINE INTERACTION. Robert M. Curtice, Victor Rosenberg, Lehigh University, Bethlehem, Pa., U. S. A.

It has been repeatedly shown that document retrieval is as much an intellectual process as a mechanical one. By combining the intellect of the user with the file searching and display capabilities of the computer, an associative retrieval strategy has been designed which enables the user personally to guide the search within the parameters of his own interests and the contents of the collection. The system eliminates the need for a thesaurus and may be used with most types of indexing. The user approaches the system with a request containing at least one term used as an index term in the collection. The term or terms are coordinated and the file searched. A list of associated terms (terms occurring together in documents) is then displayed to the user along with the number of documents initially retrieved. The term associations are determined statistically. On the basis of the number of documents initially retrieved, the user may either expand or narrow his search. From the list of terms the user selects those which seem to suit his needs and which will yield a reasonable number of documents, based on the predetermined maximum number of resulting documents. The additional selected terms are then coordinated in a manner which depends on the choice to expand or narrow the search. The user may repeat the above process with the additional terms as many times as necessary. When the user elects to conclude the search, an ordered list of documents is produced. The above type of retrieval strategy is particularly applicable to existing collections and could be optionally implemented on time sharing computer configurations. Results obtained to date on a 1,000 document collection have been particularly encouraging from the point of view of user satisfaction as well as more objective systems criteria such as computer running time and number of documents resulting from a given search.

LA DIFFUSION DES PUBLICATIONS AMERICAINES EN EUROPE AU XIXe SIECLE PAR LES ECHANGES. Juliette L. Dargent. Service belge des échanges internationaux. 16 rue des Deux-Eglises, Brux.4, Belgique.

Les publications officielles et savantes de 17 Etats d'Amérique et du Gouvernement fédéral arrivèrent dans les ministères, académies et sociétés savantes européennes par l'intermédiaire de l'Agence européenne de Paris. Alexandre Vattemare, le fondateur et directeur de cet office était l'agent des Etats américains. Cet apôtre de la coopération intellectuelle Etats-Unis/Europe apporta une conception nouvelle de promotion mondiale de la culture et convainquit les membres du Congrès de Washington d'accepter le legs Smithson, de fonder la Smithsonian Institution en lui confiant la tâche des échanges internationaux. Les passages de Wattemare dans les Etats américains sont jalonnés de fondation de bibliothèques, sociétés culturelles dont quelques unes existent encore comme les bibliothèques de Boston et de New York. Les rapports des gouverneurs des Etats fournissent des renseignements précieux sur la répartition des publications européennes arrivées grâce à Vattemare.
La documentation échangée en un quart de siècle dépasse les 300.000 volumes, enrichissement extraordinaire pour l'époque : 1840-1865.

UN ENSEIGNEMENT NON DIRECTIF DE DOCUMENTATION EN LANGUE FRANÇAISE. Jean de Laclemandiere et Françoise Gestac. U.F.O.D., 65 Rue de Richelieu, Paris (2ème), France

En matière d'enseignement l'UFOD possède une pratique qui remonte à 1946. Après 20 ans d'expérience, elle a choisi, en novembre dernier, de livrer à ses auditeurs et élèves, un cours intégralement rédigé. Les avantages de cette solution sont les suivantes: a) possibilité de transmettre une quantité importante d'information dans un temps relativement court, avec le minimum de distorsion: b) possibilité d'apporter les extensions et les additions compatibles avec les développements de la science et des techniques: c) possibilité de contact avec des auditeurs parisiens, provinciaux et francophones résidant à l'étranger: d)possibilité de dialogue intense avec et entre les élèves au cours de séances de travail fonctionnant d'après le principe d'"autogestion". Des expériences analogues dans divers domaines ont été faites en France dans le cadre de plusieurs établissements d'enseignement. Ecole Nationale Supérieure des Mines de Nancy, Faculté des Sciences Economiques de Lyon...

L'enseignement porte sur les fonctions documentaires (élection, identification, analyse de contenu, diffusion et stockage des documents), mais est considérablement élargi en direction des techniques d'information et de communication. En 60 leçons, totalisant près de 6000 pages de texte multigraphié, des centaines de tableaux, de données, de schémas explicatifs, de références bibliographiques, les points suivants sont traités: Caractéristiques de notre temps - Contraintes de l'information-Concept d'information Communication dans les groupes - Création et recherche - Chartes de l'informateur et de l'informé Image-Lecture-Notion de document-Classement-Mise en mémoire - Analyse pertinente-Synthèse des données-Linguistique-Traduction-Communication provisoire-Communication permanente- Etude opérationnelle des besoins des usagers-Classification-Sélection-Rangement-Conservation-Ergonomie-Structures documentaires-Sources documentaires - Critères d'efficience documentaire, et pondérés de "thèmes de réflexion" qui constituent les fructueux éléments de dialogue établis entre les élèves par suivent les "Cours par correspondance" et la direction de l'enseignement.

RESEARCH IMPLICIT IN INFORMATION INPUT OF NATURAL LANGUAGE TEXT. Al Delucia, John Falaimo, Rome Air Development Center, Griffiss Air Force Base, Rome, New York, U.S.A.

The increased effort in machine processing of natural language information has served to emphasize the complexity of the problem of processing text and the graphic and symbolic data associated with it. The perishability of scientific data requires the rapid processing and reproduction of whole text. Current state-of-the-art requires the separation and subsequent manual recomposition of this material after the alpha-numeric data has been processed. Experience with machine translation and machine extracting has quantified the seriousness of this problem in an operational situation. The inability to process whole text has resulted in an incomplete product and has imposed some of the most serious delays in the processing cycle. Existing machine technology has advanced to the point where it can resolve only parts of the total problem. A survey of the state-of-the-art indicates that the following techniques offer varying levels of processing capability: paper tape input typewriters; stenotype-computer processing; optical page readers, Rheeder symbolic encoding device; mechanical photocomposers; and lexical-graphical composer printers. The limitations of these techniques are considered in terms of operational applications. Although equipment in the process of development will provide significant gains in this area further development remains to be done before more general solutions to this critical area can be realized. Applied research is required to place in the inventory a spectrum of techniques for processing language information with different levels of complexity. Further innovations should be made in the areas of machine editing; page format recognition; selective photo copiers; computer processing of related graphic-symbolic information; and economical man-machine balance within the processing cycles.

UN NOUVEL ENSEMBLE DE TRAITEMENT D'INFORMATION LEXICALE. LA MACHINE BBL. Jean de Laclemandiere, Roger Bernard, Jacques Gugenheim. CERDA,50 Faubourg Poissonnière, Paris (9ème, ,France.

La machine brevetée dans plusieurs pays et commercialisée sous la marque "BBL" constitue un exemple de traitement d'information lexicale composé des éléments suivants: 1) un dispositif d'entrée: mise en mémoire de la référence ou des documents eux mêmes et de leur contenu informationnel sous une forme codifiée de manière univoque: 2) un dispositif de défilement et de mise en corrélation optique électronique de la mémoire primaire ainsi constituée et des cartes questions formulées de manière codifiée identique: 3) un dispositif de sortie: mémoire secondaire contenant les seuls éléments de réponse pertinents, sur émulsion sensible, film ou papier photographique. La mise en mémoire des concepts susceptibles de rendre compte du contenu informationnel des documents n'est pas limitée à un nombre de concepts donnés, celui-ci peut varier de quelques unités à plusieurs centaines d'unités. Le nombre de concepts mémorisables pour un document est pratiquement illimité. La mise en corrélation dynamique du contenu des documents mémorisés et des cartes questions à raison d'une vitesse de défilement de la mémoire primaire de 3 m/sec., environ 720 000 documents mémorisés par heure, est assurée dans la machine présentée "BBL 20", par le jeu de 20 descripteurs ou conditions syntaxiques combinés de manière quasi simultanée et non pas successive. La "BBL 20" comporte 20 photomultiplicateurs, des machines différentes pourraient en compter moins ou davantage. L'organe de sortie, mémoire secondaire, flash électronique est susceptible de restituer, dans le cas d'une question insuffisamment précise à l'origine, ou dans celui d'établissement de sous-mémoires partielles, 200 réponses sur émulsion sensible en une seconde. L'ensemble BBL, qui ne repose pas sur un principe de fonctionnement binaire, possède les avantages de fonctionnement suivants: 1) capacité de mise en mémoire infinie fractionnable en bobine de n' mètres, correspondant à des mémoires de chronologie, et de spécialisations diverses. 2) capacité de mise en mémoire correspondant à des documents classés uniquement en fonction de concepts décrivant leur contenu.

AUTOMATION OF INFORMATION. John R. Dorr, Consultant, P.O. Box 2845, Washington, D. C. 20013.

Automation presents specific requirements and offers new possibilities for better utilization of information. At the present time the compatibility of input-output media, languages, coding, equipment and programs is very desirable. New systems may provide this compatibility through built-in features. The direct access to precise data, selectively correlated or referring to precoordination patterns, will magnify the use which can be made of information. In order to obtain these results, changes in indexing, abstracting, and classification techniques are considered. New reference tools in machinable form - for interphase operations - will expedite the implementation of more automation applications. Format compatibility can be supplemented by multiplication and correlation of formats.

Available systems and techniques, based on educational levels and problem solving habits of the past, can be adapted to the new requirements, but new approaches can be considered and directed toward:

1. Specificity in terminology by the increase in vocabulary and the functional grouping of the vocabulary entries.
2. The use of interpretative and multifaceted correlation techniques.
3. The development of interrelated classification systems, structured for updating at the rate of growth of modern technology.

The work pursued involves the definition of these requirements and the outline for adequate solutions for the automation of information.

LIBRARIANSHIP AND THE SCIENCE OF INFORMATION.
Joseph C. Donohue, The RAND Corporation, Santa Monica, California, U. S. A.

Utilizing query responses from accredited library schools, courses offered have been analyzed to determine emphasis on various subjects, with special reference to Information Science (IS) i.e., the study of the information processes, and Science Information (SI) - the bibliography of science, operation of scientific information services, etc. Generally, schools continue to stress subjects related to operation of public and school libraries. Certain library schools are developing programs and/or courses in IS and SI, and in some cases the two are combined. Many schools now offer some training in non-traditional techniques. Increased stress on science information and the addition of courses in the new information technology have not radically altered the theoretical structure of library education. But the Information Science approach promises to contribute greatly to that structure. Librarianship possesses a long-established corpus of knowledge relating to a science of information. Librarians, traditionally service-oriented rather than research-oriented, have not exploited that body of knowledge for general principles of IS. Research should be directed to the areas of

1 Cataloging and classification- the logical and epistemological underpinnings of respective systems; the relationships of given systems to prevailing theory of knowledge.

2 The technique of reference service- for its rich potential contribution to the theory of problem solving.

3 The understanding of the social and institutional framework of the information community. Library schools need now to accompany technical instruction with research into these and similar problems in order to contribute to the theory of IS and to gain from IS better technique for the practice of librarianship.

THE RECENT MANILA SCIENCE INFORMATION SYMPOSIUM - ITS NATIONAL AND INTERNATIONAL IMPLICATIONS. Quintin A. Cala, National Institute of Science & Technology, Philippines.

The Symposium was divided into four technical sessions, namely: 1) science information communication. Fourteen papers were presented by scientists, who underscored jointly with the assembly the contentions that: i) scientific research involves communicating knowledge or facts essentially thru written technical reports; ii) technical reports should be published only when there are new significant discoveries, in simple concise terms for understanding or intelligibility to the audience for which they are intended, and as per internationally accepted printing style; and iii) a rich, dynamic library and alert, up-to-date documentation services are indispensable; 2) Information resources and their accessibility. Twenty-nine university/college, 12 government agency, and 1 industrial firm libraries reported. The university/college libraries are awfully deficient in science research references, the government agency and industrial firm libraries are comparatively well off. There is a strong desire for coordinated acquisition and a national interlibrary lending system; 3) Services to science researchers and scholars. Open-shelf service system, nationally coordinated abstracting service, adequate translation manpower, and international science information exchange system were strongly advocated; and 4) Training in science information control. There was general acceptance of the necessity of scientists taking lessons in bibliography and foreign languages, as well as of librarians taking science courses and language training. Inclusion in college curricula of science subjects and foreign languages for students aiming to be science librarians, and of library science and foreign languages for science students are deemed urgent. The national and international implications of the proceedings in scientific research work and documentation services are more meaty scientific papers in simple and clear language, simplified and more efficient documentation work, increased volume and rate of flow of knowledge on national and international level, and savings in time, funds and efforts.

INFORMATION AND DATA RETRIEVAL SYSTEMS FOR ANALYTICAL CHEMISTS.
Freeman H. Dyke, Jr., Jonker Business Machines, Inc., 26 N. Summit Avenue, Gaithersburg, Maryland, U.S.A.

Each group of scientists has a unique and distinct need for information and technical data which is not met by available information services. In many cases, the scientists band together to create their own information and data retrieval systems or encourage others to create these systems for them. This paper describes the information and data retrieval needs of the analytical chemists and reviews some of the special information and data retrieval systems created to fulfill these needs. The systems described in detail are the Joint Committee on X-ray Powder Data, the E-13 Subcommittee of American Society for Testing Materials on Absorption Spectroscopy, and Preston Technical Abstracts Services for gas chromatograph data and nuclear magnetic resonance data.

NOTIFICATION THEORY. Robert A. Fairthorne, Herner and Co., 2431 K Street, N. W., Washington, D. C., U. S. A.

Too often the proper targets of documentation are obscured by elaboration of unexamined assumptions. For instance, libraries are confused with laboratories, readers with library facilities, factual statements with facts, being informed about a document with being informed by it. Many such confusions arise from implicit belief that "Information" is a unique self-subsistent essence, like Phlogiston. We need to find which personal uses of documentary facilities can be delegated reasonably to documentary procedures. The essential first step is to examine the sparsest level of genuinely documentary activity. This still must consider source, destination, and designation, as well as the "signalling" components of message, channel, and code that suffice for telegraphist or printer. Thus we have four distinct "informations", about sources, destinations, and designations as well as that about which message has been selected. This lowest level of documentation I call "Notification". It includes Information Retrieval and Dissemination. For both these deal with selection of messages (e.g. documents) through incomplete data about source (e.g. author), destination (type of reader) and designation (topic) with given channel and code (medium) and writing system). For local, restricted, and stable interests Notification activities are amenable to formal treatments that genuinely reflect the documentary situation. In general this deliberately austere approach clarifies some existing confusions about the proper nature and aims of documentation.

(Research sponsored by the Information Sciences Directorate, U. S. Air Force Office of Scientific Research, under Contract No. AF 49(638)1427.)

INDEXER REQUIREMENTS FOR THE RECOGNITION OF SCIENTIFIC CONTENT AND CONTEXT. W. R. Foster, Associate Director, Life Sciences Division and David F. Hersey, Deputy Director, Science Information Exchange - Smithsonian Institution, 1730 M Street, N. W., Washington, D. C., U. S. A.

It is essential for the sake of economy that an information system use the least expensive personnel consistent with effective retrieval. The Science Information Exchange has for 15 years provided information concerning current research to both scientist administrators and individual investigators, with emphasis on recognition and preservation of scientific context and current research concepts. An experiment was designed to determine the extent to which subprofessional personnel with various levels of educational background can handle different types of subject content in the indexing process, with respect to efficiency as well as quality. Using a standard of correct index points determined by senior professionals (Branch Chiefs) in the light of established working definitions of topics, subprofessional personnel were able to index an average of 30% of the possible number of correct index points, as compared with 86% for the earlier professional analysis of the same projects. Subprofessional personnel took longer per topic indexed, making the estimated cost about the same for both groups. The correct index points which were missed by the subprofessional group were much more likely to be major or critical ones than in the case of the original analysts There is a strong positive correlation by rank order between educational level and performance. It should not be concluded, however, that subprofessional personnel should not index scientific research content: all of the test group did relatively well on "list" types of indexes, such as chemical substances and biological organisms. Many other types of categories were successfully handled by the test group, depending primarily upon level of educational background and experience.

EMERGING PATTERNS OF NATIONAL INFORMATION SERVICES. Bernard M. Fry, Clearinghouse for Federal Scientific and Technical Information, 5285 Port Royal Road, Springfield, Virginia, 22151, U. S. A.

National technical information requirements and the principal problems and issues involved in fulfilling them are discussed and recommendations made for steps toward their solution. The following central problem areas are examined in relation to their impact upon the structure and effectiveness of national information services in science and technology 1. The disparate though inter-related roles of data, information analysis, and information dissemination centers--together with research libraries--considered in terms of their distinct functions and their contributions to advanced information systems. 2. Achievement of network capability utilizing special information sources both Government and non-Government. 3. Development of comprehensive and non-duplicative information centers performing indexing and abstracting services for the major disciplinary fields. 4. Establishment of criteria for centralized and decentralized services where appropriate, based on effectiveness, efficiency and economy. 5. Provision of information services in various forms adapted to the needs and interests of particular scientific, technical and industrial groups. 6. Effective application of scientific and technical information to economic growth, with particular reference to the transfer process. 7. Ways and means to obtain user feedback and to measure the effectiveness of information services.

THE USE OF CURRENT SCIENTIFIC RESEARCH INFORMATION FOR ADMINISTRATIVE PURPOSES. Monroe E. Freeman, Director, Science Information Exchange - Smithsonian Institution, 1730 M Street, N. W., Washington, D. C., U. S. A.

Information concerning who is doing what in current research is useful to agencies administering grant programs, and the Science Information Exchange of the Smithsonian Institution is primarily oriented to this service. Other directly related purposes of the Exchange include the provision of research information to laboratories sponsored by Federal grants and contracts, and to individual recipients of these grants.

It is of continuing interest to the Exchange to evaluate the types of scientific subject matter information most often called for by these three groups of users. Data collected to date suggest that the first class of users requires the full spectrum of S.I.E. facilities, both general and detailed information on subject questions, usually associated with some administrative parameters as well. The need for broad surveys of subject areas is greater for this group, as might be expected; there was a surprising frequency of need for detailed breakouts of these broad areas on the part of grants administrators. The second and third classes of users require primarily the more specific types of subject matter coverage, although many programs in the second class do require rather generic level searches.

More recent and complete data will be presented on the points just discussed, and the relationship of services to the individual scientist to those of a more administrative orientation will be considered.

A NEW TYPE OF CONCERN IN TERMS OF DOCUMENTATION: SCIENTIFIC AND TECHNICO-COMMERCIAL DATA PROCESSING. Sylvia Gabbai, Euratom, European Atomic Energy Commission, Ispra, Varese, Italy.

The Documentation Unit of Ispra's Euratom Technical Services takes care of a very peculiar kind of documents which consists of catalogues, pamphlets, leaflets and research studies sent by research or production firms for the use of Scientists. This Scientific and Technico-Commercial Documentation became so essential for the use of all technicians within Ispra's European Atomic Energy Commission Research Center, that a special classification has been worked out and data is stored by Flexowriter for the mechanization of the system through I3M 1401.

FONCTIONS ET ACTIVITES DU GROUPE D'ETUDE SUR L'INFORMATION SCIENTIFIQUE : "DOCUMENTATION DE LA DOCUMENTATION". Natacha Gardin, Section d'Automatique Documentaire, C.N.R.S., Marseille, France.

Le Groupe d'Etude sur l'Information Scientifique a été créé en 1963 par la Délégation Générale à la Recherche Scientifique et Technique. Ce Groupe est chargé d'établir un "inventaire des études et des applications en cours intéressant le traitement automatique de l'information scientifique, principalement la documentation et la traduction automatique". Cet inventaire se matérialise sous la forme d'un fichier analytique établi à partir de la documentation technique non publiée et des revues spécialisées dans les domaines en question. A ce jour, 5000 documents ont été dépouillés (à dater de 1961). Le fichier comprend d'une part un ensemble de rubriques correspondant aux différents aspects du domaine, et d'autre part une série de subdivisions dites "annexes", concernant les Chercheurs, les Centres de Documentation, les Congrès, les Equipements, etc. Un "Lexique de Descripteurs" permet l'analyse de ces textes et ses différents chapitres donnent le plan du fichier. Ce lexique, bilingue pour l'instant, (anglais - français), sera quadrilingue dans sa version définitive (+ russe - allemand). Auparavant, l'on espère parvenir, avec la collaboration d'organismes étrangers homologues à une uniformisation des principales notions qui permettent d'indexer la "Documentation de la Documentation". Enfin le Groupe a mis au point une enquête sur les Moyens Nouveaux en Documentation Scientifique et Technique en France, enquête équivalente aux "Systèmes non-conventionnels d'information technique actuellement en usage" publiés par l'Office of Science Information Service de la N.S.F. L'ensemble des travaux du groupe permet ainsi la fabrication d'ouvrages de références, de bibliographies spécialisées à partir du fichier analytique, en même temps qu'il va dans le sens des recommandations de la F.I.D. en matière de coopération internationale en Information Scientifique.

APPLICATION OF ANALYSES OF TOTAL DEMAND FOR SCIENTIFIC AND TECHNICAL INFORMATION. C. [illegible] Hanson, Aslib, 3 Belgrave Square, London, S.W.1., England.

A basic aim common to two separate but allied investigations (A and B) was to reveal the patterns of demand for scientific and technical information, and to examine the demands for various types of information against the provisions for meeting them. The specific aim of investigation A was to study scientists and engineers in action over long periods. Five research workers were given small tape-recorders. They recorded daily, for periods up to twelve months, their problems and actions, with particular reference to their sources of information and ideas, both written, formal and informal, oral as well as printed. The specific aim of investigation B was to obtain details of every visit of library users which took place in a technical library over a given test [illegible]. The data were gathered by questionnaire supported by observation. The results of the two investigations taken together, examined in conjunction with earlier results published by ourselves and other workers, yield evidence that the total demand for information is multiform and that the provision for meeting it must likewise be multiform: no one system of making information available can meet adequately more than part of the total demand. That, of the total [illegible], about one fifth was for [illegible] future or [illegible] fact, [illegible] half for information readily [illegible] in a working or subject library. About [illegible] was for [illegible] the enquirer's central subject field. There was a [illegible] for comprehensive searches of [illegible] of literature; and [illegible] specific occasions [illegible] was sought for information about commercial [illegible] and [illegible]. Appropriate quantitative measures of these [illegible] in terms of need are given as [illegible] of the [illegible].

DESIGN DECISIONS FOR AN INTEGRATED LIBRARY SYSTEM. Marjorie Griffin, International Business Machines Corporation, Advanced Systems Development Division, Los Gatos Laboratory, Los Gatos, California, U.S.A.

The library's responsibility in documentation begins with the entry of the bibliographic information which is the source of all subsequent library records for each document acquired. The steps from acquisition to circulation are now mechanized in many libraries, but seldom as part of an integrated system. Such a system has been developed in an experimental technical library ideally situated: it is a working library for an IBM laboratory and serves as a testing ground for advanced data processing equipment and programs. To keep up with the increasing information supply and demand, this library was moved from piecemeal mechanization to an integrated system which now provides mechanized bibliographic control and will eventually include mechanized information searching and retrieval. Design decisions for this system were based on an analysis of the effectiveness of existing procedures as evaluated by both the library staff and library users. The basic question was not "How can we mechanize this operation?" but "Why do we do this...and why do it this way?" The success of the system now implemented depends directly on the thoroughness of that questioning. As the answers disclosed duplicate effort, avoidable sources of error, and unfilled needs, the system design was modified as permitted by the equipment to be employed. In operation, a single machine-readable entry establishes bibliographic control when a document is ordered, and on-line processing maintains this control, automatically accommodating corrections and revisions as needed. The same input, printed out selectively, provides purchase orders and subject catalogs. With the flexibility demonstrated and the unexpected value of unplanned byproducts, this integrated system seems to successfully anticipate both the future needs of the library and the possibilities of present data processing technology as it applies to the phases of documentation which are concentrated in the library.

A RETRIEVAL ALGORITHM PERMITTING RANKING OF DOCUMENTS RETRIEVED BY DEGREE OF PERTINENCE TO THE USER. Mary K. Hawes, UNIVAC Div. of Sperry Rand Corp., 650 N. Sepulveda Blvd., El Segundo, Calif., U.S.A.

A Retrieval Algorithm has been developed to meet the requirments for an on-line document retrieval system that can be expanded to an information retrieval system. These requirements are: (1) relatively complex requests can be formulated in a straightforward manner without necessitating assistance from a retrieval specialist, (2) the language and the rules used to specify the criteria permit the user to state both minimal requirements and maximum pertinence, (3) documents retrieved can be ranked by degree of pertinence to the user, (4) the request statement is suitable for direct entry into a computer system, (5) the request statement can be verified as being processable while the requestor is on-line to the computer system, (6) procedures and machine coding for processing the request can be generated dynamically, (7) procedures allow processing of multiple requests in a parallel operation, (8) techniques used for demand requests are suitable for information dissemination, (9) the retrieval algorithm can be expanded for information retrieval.

In order to meet the above requirements, it was necessary to develop some new techniques. Among these are Information Grouping Logic (IGL), an extension of Boolean logic incorporating the minimal requirement concept, a problem oriented language for stating the retrieval request, and a diagrammatic representation of the request. Decision Tables are used to define the Algorithm which generates the procedures for processing the request and ranking the documents retrieved by degree of pertinence to the user. A number of examples are included.

CATALOGING AND RETRIEVAL STATUS OF STATE EXPERIMENT STATION TECHNICAL BULLETINS AT INSTITUTIONAL LIBRARIES.* John F. Heer, Publications Office, Iowa State University, Ames, Iowa, U.S.A.

The volume of scientific-technical literature has increased substantially, and traditional respository libraries have had growing problems of classification, cataloging, storage capacity, financing and staffing. It seemed probable that these problems had necessitated some changes in priorities for the collection, cataloging and binding of serial technical bulletins of the state experiment stations. By mail questionnaire, an exploratory survey was made of the general and subordinate libraries of land-grant colleges and universities concerning the disposition of three series of Iowa State University publications. For each series, libraries were asked (a) if the series was cataloged, (b) if it was preserved in the permanent collection, (c) what disposition was made if not cataloged and preserved and (d) to indicate why a series was not cataloged or preserved. The returns showed that, with few exceptions, the institutional libraries surveyed were cataloging and preserving the three series but that some changes in priorities and handling had taken place. Fewer libraries than formerly are cataloging individual bulletins, and more are cataloging by series only, keeping each cataloged series accessible "in one place." Several libraries indicated being in arrears in binding a series because of binding-budget limitations or because experiment station publications had been forced into lower priorities than formerly. (Volume contents and indexes are furnished to libraries upon completion of volumes in two of the three Iowa series.) Cataloging of the bulletins by series rather than individually is not too serious from the standpoint of retrieval if each series is accessible and bound with volume contents and indexes. Without other specific documentation and indexing, however, lack of both individual cataloging and volume binding is potentially more serious since usefulness of the volume contents and indexes as retrieval aids also may be reduced or lost. (*Journal Paper No. J-5078 of the Iowa Agricultural and Home Economics Experiment Station, Ames, Iowa. Project No. 102.)

COMPARISON OF KEYWORD INDEXING AND INDEXING BY SYSTEMATIC CLASSIFICATION. Kenzo Hirayama, Fuji Iho* F.L. Research Laboratories, Kinbanashigara, [illegible], Kanagawa, Japan.

There are two following ways of indexing the concepts contained in a literature in order to retrieve the literary information from its content.

1) Indexing by systematic classification [illegible] concepts are classified systematically beforehand, and [illegible] to which classified item belongs the concept [illegible] article. Dewey Decimal Classification and Universal Decimal Classification (U.D.C.) are examples of such classification.

(2) Indexing the keywords [illegible] This system is similar to the subject indexing of a book. [illegible] symbols) that seem useful as index are used as index words. In some cases, index words are [illegible] standardize from the point of synonym. The index, [illegible] descriptors index belong to this [illegible].

In order to clarify the [illegible] and disadvantages of these two methods, [illegible] compared with actual examples. One of the [illegible] patents and examinations were [illegible] a commercial index [illegible] patents were indexed by keywords, and how the patent was classified and indexed by U.D.C.

Indexed keywords { Words that were appropriate as index........7 / Words that were inappropriate as index.....21 }

Appropriate keywords not indexed but should have been indexed from comparison with the indexed keywords..........16

In U.D.C. system, all the concepts represented by the foregoing appropriate index words can be expressed by three classification numbers. In addition, some examinations were also made on the time and labor required for classification and noise found in retrieval. It was thereby found that it is easy to pick out the keywords but quite difficult to make this uniform, and this results in loss in retrieval and too much noise. In contrast, systematic classification was found to take more time but gave better retrieval rate. Consequently, the two systems should not be considered as interchangeable but should be used as complementing each other.

UNITERM AND THE LEARNING CURVE. Henry G. Higley, Department of Research and Statistics, American Chiropractic Association, 920 East Broadway, Glendale, California, U.S.A.

The documents that are incorporated into the collection being developed by the Department of Research and Statistics of the American Chiropractic Association consists mainly of reports of studies dealing with the human spine. All documents are classified and coded according to concepts. We use a uniterm system of coordinated indexing. Uniterm cards were classified by the number of accession numbers per card. Five hundred documents representing a three month imput resulted in 258 uniterm cards with 27?5 accession numbers. The number of accession numbers per uniterm card ranged from 1 to 159. The frequency of uniterm cards with a given number of entries decreased as the number of accession numbers increased and closely followed the curve $y = ax^b$ with the exception of $x = 1$. The study of the imput over a three month period of other uniterm systems leads to similar results. It is suggested that the parameters a and b represent characteristics of the literature and of the system, and that $x = 1$ represents new concepts and older concepts being replaced by new ideas. The equation $y = ax^b$ often referred to as the learning curve or as the progress curve has been studied by many and it has been used for cost estimating and for studies of production progress. Houston and Wall in a paper, The Distribution of Term Usage in Manipulative Indexes, American Documentation, Vol. 5 (1964) pp 105-114, discussed the cumulative distribution of terms in the index vocabulary of 10 collections of various sizes. The distributions are made linear by log-normal transformations. Our study deals with imputs over a short period of time and deviations from the prescribed curve are readily detected since they are not diluted by the whole collection. A significant deviation in the frequency distribution of terms used in a short interval imput may be a warning sign to look for some assignable cause such as a tendency on the part of a document analyst to incorporate concepts into terms already in the vocabulary list, even if the relationship is remote.

TOWARD AN INTERNATIONAL SOCIAL WELFARE DOCUMENTS SYSTEM. Joe R. Hoffer, International Conference of Social Work, 345 E. 46th Street, New York, N. Y., U.S.A.

Documenting and exchanging knowledge is an inseparable part of planning, research and development in any field of endeavor - and social welfare is no exception. The problem of creating improved methods for searching and retrieving data is difficult because of the nature and complexity of social welfare. Social welfare is defined as a "field" encompassing the community social services under governmental or non-governmental auspices which exist potentially for each member of the community. It is clear that the future international use of social welfare knowledge will depend upon the development of a simple inter-communicating network of cooperating specialized centers in major information producing countries. It is proposed that a network of these document centers be organized. This network will be a federative, non-centralized organization that will require a minimum of central services and a maximum of autonomy in the existing information centers. The basic elements for the proposed coordination of international documentary work in social welfare include: adoption of a system of index terms; exchange of machine readable punch paper tape containing invariant bibliographic data, annotations, and/or abstracts; publication of machine-produced indexes and selected bibliographies, and provision of copies of documents requested. Each element is discussed and concrete proposals are presented based on the recommendations of three Workshops and several pilot studies conducted by the United States National Conference on Social Welfare. Implications for international exchange of knowledge are discussed with specific reference to communication problems faced by the International Conference of Social Work in the conduct of twelve international conferences held on four continents.

INDEXER CONSISTENCY TESTS - ORIGIN, MEASUREMENTS, RESULTS AND UTILIZATION. Robert S. Hooper, International Business Machines, Federal Systems Division, 7220 Wisconsin Ave., Bethesda, Maryland, U.S.A.

The efficiency of a document reference retrieval system may be measured in terms of the success in retrieving documents. This success is often measured in terms of "recall" and "relevancy" ratios. Efforts are extended to make these ratios as large as practical to meet the needs of a real system. Recall ratios below the 100% mark are often attributable to indexing failures. The nature of indexing failures can be determined by studying why documents are not recalled, if it can be assumed that the collection is sufficiently large and sufficiently well-known to enable the measurement of recall, and if enough search requests are available for study. Another method of determining these failures, is to perform a controlled indexer consistency test where indexers will, each in turn, index a significant number of "live" documents. By analyzing 17 reports of indexer consistency tests, it has been found that there is no standard measure of consistency; hence, a set of three measures has been derived and is used to analyze the data in these reports. The analysis indicates that a range of consistency values between 10% and 80%+ may be expected, depending upon the indexing parameters. Additional analysis shows that uniquely structured classification schedules and indexer training programs are the indexing parameters most often associated with the higher consistency results. Further, the analysis suggests that consistency tests have an important role as a tool in the design of indexing systems and in the training of indexers and quality control of the indexing.

SUBJECT INDEXING USING A COMPUTER-MANIPULATED THESAURUS AS VOCABULARY CONTROL. Marjorie R. Hyslop, American Society for Metals, Metals Park, Ohio, 44073, U. S. A.

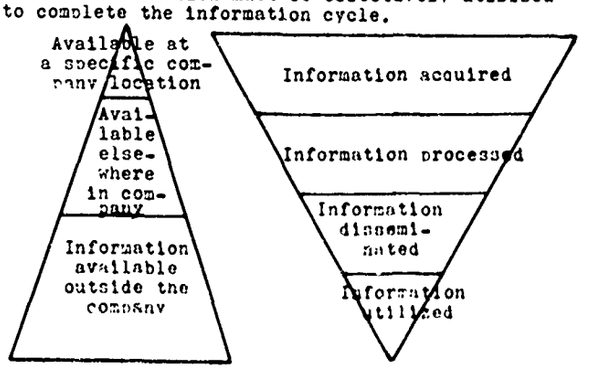
In redesigning the information retrieval system of the American Society for Metals, it was decided that the capability should be included to automatically generate a subject index to the Review of Metal Literature (abstract journal). It should be generated as a computer product of the over-all information retrieval system, and based on a thesaurus vocabulary. Previously the index was manually prepared utilizing a modified classification scheme as authority. The first step was to compare the two types of vocabularies. For test purposes the "EJC Thesaurus of Engineering Terms" was used, recognizing that it is more restricted than will be the eventual ASM Thesaurus. Nevertheless, comparative tests on 100 abstracts showed that 83% of the indexing concepts could be intellectually duplicated by the thesaurus vocabulary. Tests for indexing consistency were performed by comparing entries submitted by 9 indexers for the same body of 25 abstracts. Indexing rules were devised based on these tests and also on the thesaurus structure and the computer processing capabilities. One of the important features of the new method is the merging of two intellectual indexing functions previously performed separately on the same document--deep indexing for machine retrieval and shallower indexing for publication. Other features are (1) the provision of two levels of indexing, (2) the automatic generation of cross references, (3) the provision of monthly indexes in addition to a cumulated annual index, and (4) the inclusion of titles with index entries.

FORMATION SUPERIEURE ET DOCUMENTATION. Jean-E. Humblet, Bruxelles, Belgique et Nice, France.

L'auteur a étudié, en Belgique, en France, aux Pays-Bas et au Royaume Uni, quels cours dans des Facultés universitaires de Sciences appliquées (technologie, médecine, etc), de Sciences sociales (droit, économie, sociologie, etc) ont pour objet les recherches et les techniques documentaires. Il distingue les besoins d'initiation de tout universitaire aux méthodes scientifiques et à la recherche documentaire comme utilisateur de bibliothèques et services de documentation, et la formation spécifique des spécialistes de la documentation. L'auteur insiste sur l'aspect épistémologique du problème de l'information comme élément fondamental du processus d'acquisition du savoir au niveau universitaire. (choix et utilisation des documents, sélectivité et exhaustivité). Il décrit ensuite sur la base des travaux de JACOB (France), de l'American chemical Society, de SHERA et KENT (Cleveland) les diverses fonctions documentaires de l'assistant-bibliothécaire au chef du service de documentation, du lecteur occasionnel à l'information scientist. L'auteur étudie les programmes d'enseignement idéaux, d'une part en vue de sensibiliser aux problèmes documentaires tous les futurs chercheurs et cadres et de créer des ponts entre ces hommes et les documentalistes et "documenticiens", d'autre part en vue de former des spécialistes. En conclusion, les suggestions suivantes sont formulées : a) organiser un colloque paritaire (professeurs d'universités et spécialistes de la documentation) afin de mettre au point un programme et des méthodes pour initier à la documentation, compte tenu des éléments généraux et de ceux qui sont spécifiques aux diverses disciplines, b) préparer un programme de formation à chaque fonction énumérée, c) prévoir une partie de la formation sous forme de colloques et séminaires, au niveau international (régional) en effet, a) la dimension nationale est souvent trop réduite pour la formation de certains spécialistes, b) le progrès des techniques documentaires postule des solutions internationales.

THE INFORMATION SERVICES OF A TECHNICALLY-ORIENTATED FIRM -- PRINCIPLES FOR ORGANIZATION. Eugene B. Jackson, International Business Machines Corporation, Armonk, N. Y., U. S. A.

The following eight principles have evolved from association with a variety of industrial information activities and are believed to have general applicability: I -- Information is a vital asset of the firm. II -- Information must be totally exploited in support of the company mission. III -- The company information complex must be a "good citizen" of the total information community. IV -- Information is an important company "product" V -- Company information systems must acquire all essential information. VI -- Information acquired must be processed so as to be under effective control. VII -- Information on existence of this newly available body of information must be properly disseminated. VIII -- Information must be effectively utilized to complete the information cycle.



GENERALIZED INDUSTRIAL INFORMATION "TRIANGLES"

EXTENDED WADEX SYSTEM: TOOL FOR BROWSING, SEARCHING & EXPRESS INFORMATION WITH ADJUSTABLE INTELLECTUAL PREPARATION EFFORT. S. Juhasz, Applied Mechanics Reviews. 8500 Culebra Rd., San Antonio, Texas, H. Wooster, AFOSR, Washington, D.C., E. A. Ripperger, Univ. of Texas, Austin, D. Falconer, AMR.

WADEX (Word Author inDEX) an IBM 1401 prepared mechanical index, extension of KWIC concept. This paper describes greatly improved third version where intellectual effort ranging from zero to considerable can be employed depending on titles and purpose. WADEX system usable for browsing, searching, express information, also for preparation of specialized deep indexes of specialized holdings. First experimental WADEX (1) and second experimental WADEX (2) are based on 8,000 titles of APPLIED MECHANICS REVIEWS (AMR), an international critical review magazine. In WADEX titles printed fully with author names, and as many lines used as necessary (60 characters per line). WADEX entry: word in title (except forbidden word) or author's name. Alphabetically sequenced entries printed out of context. Words are single words or hyphenated word pairs or "Tagged Words" (TW) Latter consist of significant "word compounds" in which one or more words are forbidden words. TW selected from subject index authority list of AMR and hyphenated words from frequency list of words of previous WADEX. For more than one identical word entries, titles alphabetized according to first author and for more than one identical author entries according to reference number, which printed in line with entry. Multiple entries listed only once per column. WADEX arranged in 2 columns, both columns printed simultaneously with dictionary entries at top and pagination at bottom by computer. After addition of 26 alphabetics "Compuscript" is the photo-ready copy with 6.5 average entry per paper including 1.5 for authors. Published WADEXes based on available titles. Present paper extends system for titles with intellectually added keywords, all individually in brackets. Documentation of machine program now available. AFOSR, NSF & ONR are sponsors of WADEX and/or AMR.

Reference 1, WADEX (Word Author inDEX) A new tool in literature retrieving, E. A. Ripperger, H. Wooster, S. Juhasz: MECHANICAL ENGINEERING, March 1964, pages 45-50, Reference 2, E. A. Ripperger, H. Wooster, S. Juhasz & D. Falconer. Second experimental WADEX for AMR Vol. 16, Govt. Printing Office, July 1965, 650 pp.

ENTROPY AS A MEASURING TOOL FOR INFORMATION SYSTEMS. Jay Hilary Kelley, Rutgers, The State University, New Brunswick, N.J., U.S.A.

A true measuring tool for information systems, by necessity, must include 1) a means for measuring the needs of users and the value of the information. But it also must include 2) mechanistic aspects not dependent upon the variability of value or need. Since mechanistic aspects do depend upon the physical organization of information, they can be measured quantitatively. "Entropy of Knowledge" is proposed here as a tool for making this measurement. The use of entropy in this context is natural since the mechanistic aspects of information systems resemble the thermodynamic systems where entropy has been used effectively for decades to measure the degree of organization of apparatus. Entropy can be considered a probabilistic entity; and in both informational and thermodynamic systems probabilistic and/or statistic factors are prevailing. It is suggested that through behavioristic studies, entropy values could be shown to have a relationship to the needs of users and to the value of information. In fact, a well organized information system should allow for the maximum exercise of user's choice of information according to need and value. Since entropy values would indicate how well information is organized, they should indicate how easily a needed element of information can be retrieved. Some discussion is devoted to how one evaluates information systems using entropy concepts.

OPERATION OF THE AIR FORCE MACHINABILITY DATA CENTER, John F. Kahles, Air Force Machinability Data Center (AFMDC), 3980 Rosslyn Drive, Cincinnati, Ohio, USA.

This paper outlines the basic design concepts leading to the development of AFMDC, reflects procedural changes made since its design was completed, and discusses actual operation of the Center since it was started on 1 Oct. 1964. Strong economic justification for establishment of AFMDC is indicated in that more than $34 billion are spent annually for labor and overhead in the metalworking industries alone in the United States. Other significant economic data include expenditures for machine tools amounting to more than $1 billion annually plus $787 million for machine tool accessories, including small expendable cutting tools. Machining information is stored on a set of 10 punch cards, including general information as well as specific numerical machining data such as speeds, feeds, depth of cut, tool material and geometry, cutting fluids, and other significant variables. The entire system is based upon the "machining situation" concept. A machining situation is defined when a particular material, with definite chemical, physical, and mechanical properties, is being processed by a specific material removal operation, either a conventional or nonconventional process, and it is the focal concept for acquisition, interrogation, or presentation of information. This paper shows how forward-looking design concepts can anticipate updating of a system, such as conversion from punch card storage to computer. Computer programs have been developed and demonstrated for converting coded machining data to printout in machinist's language. The paper summarizes the inquiry experience with more than 300 specific inquirers and discusses procecures used in development and maintenance of a User File.

PROGRAM OF THE CENTER FOR INTERNATIONAL BIOMEDICAL COMMUNICATIONS RESEARCH. Allen Kent, Director, Knowledge Availability Systems Center, University of Pittsburgh, Pittsburgh, Pennsylvania 15213, U.S.A.; Pierre J. Vinken, M.D., Executive Chief Editor, Excerpta Medica Foundation, Amsterdam, Holland.

Research aimed at studying the most effective means of transmitting new medical information from research to the practicing physician will be conducted by the University of Pittsburgh-Excerpta Medica Foundation Center for International Biomedical Communications Research. The Center will investigate techniques of information retrieval, publishing and librarianship, as well as computers and other hardware elements in an attempt to develop an integrated communications system which will provide the physician with medical knowledge while it is new and vital. The joint University of Pittsburgh-Excerpta Medica Foundation project will utilize, as the chief source of research material, the Foundation's twenty-year collection of about 1,300,000 abstracts, indexes to the Excerpta Medica specialized abstracting journals, and the microfilm library in which original articles from over 3,000 biomedical serials are stored. The initial research will be concentrated in three areas: a selective dissemination program for physicians; a study of the problems relating to medical terminology; and an investigation of the feasibility of applying automatic indexing techniques to medical literature.

CITATIONS IN SOVIET METALLURGICAL LITERATURE. Ladislav Kofnovec, Institute for Technical and Economic Information, Konviktská 5, Praha 1, Czechoslovakia.

The present study examines the citations given during a year in three Soviet journals "Stal", "Metallovedeniye i termicheskaya obrabotka metallov", and "Fyzika metallov i metallovedeniye" which deal with the problems of metallurgy, heat treatment, and metal physics. The paper studies the cited documents from the viewpoint of type of document, language, country of origin, and age. The results were tabulated and graphed. The most cited journals were tabulated both for all three source journals together and for each source journal separately; the results described in the present paper were compared with the Burton's results. The cited documents were divided into two groups: up to ten years old and older than ten years. Most cited were the documents in Russian, further in English, German, and French. The documents in Russian, English, and German were cited in 97.6%. In journal citations not older than ten years Russian participates with 59-63.9%, English about 25 %, German at the most 8.1%, and French at the most 2.5%. In comparison with the older journal citations Russian and English increased their share whereas the percentage of German and French journal citations dropped. 50 % of cited papers from the last ten years appeared in 3.8-4.4 % out of the total number of journals cited. The first 30 journals most cited in all the source journals (not older than ten years) include 10 most cited journals in all three source journals separately. From the 19 most cited Western journals 16 titles can be found among the 20 most cited journals according to Burton's paper.

SEARCHING FOR SUBSETS IN MACHINE RECORDS OF CHEMICAL STRUCTURES AT THE CHEMICAL ABSTRACTS SERVICE M L Krakowsky, R W White, W C Davenport, Chemical Abstracts Service, Columbus, Ohio

The CAS computer-based substructure search system retrieves information about compounds on the basis of structural characteristics.

This paper describes the composite of techniques used in the present system. These techniques range from the use of simple screens to iterative atom by atom searching. One important phase under development employs set reduction to decrease the amount of unproductive searching.

The main emphasis of the paper is on screens by means of which inappropriate compounds are rejected, promptly and without iterative search. Criteria for evaluating screens are based on a predicted question pattern and include development costs, run times, and per cent screenout. Other factors discussed are the effect of multiple file organization, the value of nested machine records, and the relative merits of special purpose screens versus blanket screens.

Pilot results are presented and are projected to a full scale system. Present and proposed output formats are described, and the importance of interaction with users and potential users is stressed. This work is partially supported by funds from the Department of Defense, the National Institutes of Health and the National Science Foundation through a contract with the National Science Foundation.

P I C S: THE PHARMACEUTICAL INFORMATION CONTROL SYSTEM OF MERCK SHARP & DOHME RESEARCH LABORATORIES. Margaret C. Kolb, Jerome T. Maddock, and Barbara H. Weaver, Merck Sharp & Dohme Research Laboratories, Rahway, N.J., U.S.A.

The Pharmaceutical Information Control System (P I C S) developed at Merck Sharp & Dohme Research Laboratories provides centralized control and methodology for a series of decentralized information areas in the Division. It is compatible with and instrumental in total data processing and analysis of research information. Serving as a Core Index to all information resources of the Research Laboratories, it also processes, stores, and retrieves research project information for the staff members for planning and retrospective searches. A register of all domestic and international clinical research information on experimental and in-line products of Merck & Co. Inc. is provided by this system.

An eight digit dual-faceted classification code was developed based on a company-wide program identification scheme. This code of two mutually exclusive facets enables us to identify a product with a field of research. This code has been adopted for use in administrative planning, cost accounting, time allocation, and internal reporting. The uniform use of this information code within the Research Division minimizes the vocabulary barrier between the use and information system and provides the system with a self-indexing device for internal reports.

Incoming mail is copied and registered by the information center prior to transmittal to the addressee. Copies of all outgoing mail and intramural correspondence are directed to the center. An information scientist analyzes each document and selects the project code and document descriptors. This information is punched into 80-column cards. The documents are filed by code and the cards are filed alphabetically by name, term, and by date. Output forms include information, documents, printouts of document citations, and reports to drug regulatory agencies. Continuous system evaluation results in reduced I/O time, better utilization of personnel, and improved user feedback and contact.

A RETRIEVAL PROFILE FOR CURRENT RESEARCH INFORMATION. Frank J. Kreyche, Associate Director, Physical Sciences Division and W. R. Foster, Associate Director, Life Sciences Division, Science Information Exchange - Smithsonian Institution, 1730 M Street, N. W., Washington, D. C., U. S. A.

The Science Information Exchange (S.I.E.) provides information concerning details of ongoing research in the sciences to Federal and private agencies engaged in or supporting such research, as well as to individual investigators working on specific problems. Emphasis is placed on professional level retrieval of current research concepts. To assess the adequacy of such retrieval, it was desirable to have an evaluation of the retrieval profile in routine use. In recognition of the subjective nature of the determination of "relevance" of an answer, an experiment was designed in which agreement was obtained on a body of information constituting an answer to a specific subject question with a technically trained representative of a Federal agency whose program is included. In this way the recall of the S.I.E. system with respect to that agency's projects was assessed along with the relevance determination. Results to date indicate that under the conditions of the S.I.E. operation, with the same professional scientists doing both indexing and retrieval, reviewing all outgoing subject material, it is possible to have a high level of both recall and relevance. Index structure is an important contributory factor in this result, in that relevant concepts are not scattered alphabetically but are grouped by and for the use of scientist-indexers.

WHITHER PERSONAL INFORMATION STORAGE AND RETRIEVAL SYSTEMS. Robert M. Landau, American University, 1900 F. Street, N. W., Washington, D. C.

Dr. Vannevar Bush compiled a amazing set of predictions in his article, "As We May Think" in the July 1945 Atlantic Monthly. He predicted a desk size computer-driven highly miniaturized, personally organized, individual document storage and retrieval system -- an extension of one's memory; therefore, he dubbed it Memex. All of his predictions are essentially within the state of the art today. Yet we have no Memex. However, more and more people are requiring better and better organized information systems for personal and professional use. This is happening while simultaneously the amount of information to be stored is mounting at an ever increasing rate.

Large institutionalized document storage and retrieval systems are growing rapidly. Many are computer-driven. Humans often seem to be considered as peripheral to these systems. Consideration should be given to increasing the efficiency as well as reducing the cost of personally oriented information storage and retrieval systems. The recent development of multiple access time sharing systems like Project MAC and implicit programming show promise in this area.

It is proposed that an all out effort be instituted in terms of equipment, program languages and classification schemes to make available to a large number of individuals in the next decade an updated version of Memex. The problem is not one of gigantic strides to be made in the state of the art as much as it is a reorientation of our human efforts to accomplish this goal.

ORGANISATION ET CONSULTATION D'UN THESAURUS. Levéry Francis, IBM France, 168 rue de Rivoli, Paris, France.

Lorsque l'on interroge un fichier documentaire organisé à l'aide d'un système de mots-clés ou de descripteurs, il est en général nécessaire de consulter préalablement un thésaurus permettant de déterminer les termes nécessaires et suffisants pour définir la demande de documentation. La recherche de ces termes revient à déterminer le vocabulaire qui constitue l'environnement sémantique d'une notion. Cet environnement peut être automatiquement déterminé en considérant que deux notions sont d'autant plus voisines que leurs définitions sont plus semblables. Il est donc possible de considérer un thésaurus comme un fichier de documentation lexicologique concernant les mots-clés. Chaque mot-clé est inuexé à l'aide de définisseurs ce qui permet une sélection automatique de l'environnement. L'avantage essentiel de cette méthode est de permettre une construction facile du thésaurus, puisque les relations sémantiques sont déterminées automatiquement. Il est d'autre part possible de calculer des environnements plus ou moins larges, en imposant un nombre de définisseurs communs plus ou moins grand entre deux notions pour les considérer comme reliées. Les moyens automatiques et les programmes utilisés pour la sélection automatique de documentation peuvent servir à l'exploitation d'un tel thésaurus. Une application expérimentale a été réalisée dans le domaine médical. Un thésaurus a été constitué pour des noms de maladies, en utilisant un vocabulaire fondamental de définisseurs. L'expérience a montré que l'environnement déterminé automatiquement pour chaque notion était excellent et permettait de ne plus avoir de documents manquants dans les reponses.

"LE SERVICE DOCUMENTATION" COMPARE AUX BREVETS, DANS LE DOMAINE NUCLEAIRE. Laurice Lenoir, Société Brevatome, 44, Avenue du Président Kennedy PARIS 16ème - FRANCE

Pour aider la recherche, le Commissariat à l'Energie Atomique a confié à sa filiale brevets, le développement, en liaison avec l'industrie privée correspondante, d'une documentation brevets complémentaire de la documentation conventionnelle. Elle est utilisée par les chercheurs et les services de propriété industrielle (dépôts et licences). Bien qu'il existe des organismes spécialisés en brevets, que ces documents sont d'une manipulation complexe et délicate, cette initiative a été jugée indispensable :
- le brevet est l'aboutissement de la recherche et doit constituer le point de départ de la recherche à suivre. Ce jalon doit être connu avec précision. - les techniques nucléaires sont des techniques de pointe, à évolution rapide; les investissements sont élevés et les enjeux importants. - le brevet est l'information technique et technologique par excellence. Il représente à lui seul le 1/10 des publications techniques. Pour éviter toute dispersion, tout cloisonnage, il fallait disposer rapidement de cette information, dont le traitement et la diffusion ne pouvaient pas être laissés aux chercheurs, le brevet étant un document particulier, rarement attractif dans sa forme. Ainsi 7.500 brevets environ sont sélectionnés chaque année. Chacun d'eux subit les traitements bibliographiques habituels, fait l'objet de recherches de surveillance et de correspondance, puis est intégré en collection de recherche. De plus, l'information est distribuée : dans les bibliothèques des centres, aux industriels qui le désirent, et même aux chercheurs. Une revue analytique (1) est éditée, depuis un an en association avec l'Euratom, et avec l'aide de plusieurs centres nucléaires européens qui élaborent les résumés techniques des brevets (dans la langue d'origine et en anglais). Cette collaboration a permis d'éviter la répétition de travaux identiques en plusieurs points de la Communauté, et d'améliorer la qualité et la pénétration de l'information. Celle-ci est mise en mémoire sur bande perforée Friden, et peut être transposée sur carte ou bande magnétique I.B.M.

(1)"La Propriété Industrielle Nucléaire - Atomic Patent Abstracts"

ECONOMIC ANALYSIS OF A TECHNICAL INFORMATION DISSEMINATION SYSTEM. Nathan P. Levy and Ross M. Sigmon, Western Electric Company, Inc., P. O. Box 900, Princeton, N. J., U.S.A.

A crucial question posed by management when developing a technical information dissemination system is: will the participating engineers gain sufficient benefit from its service to pay for the system's operating costs? An economic study, conducted on an experimental basis in 1964 in Western Electric, indicated that a technical information system can pay for itself if it services those engineers with needs for relatively large amounts of information. This conclusion was based on quantitative estimates obtained from surveying the 121 engineers (mainly research, development and planning engineers) from eight locations, who were participants in an experimental technical information dissemination system. The questions concerning the economic evaluation were based on a concept which is described as the "Cost of Information Transfer." This paper describes this concept in detail and relates the concept to a real technical information dissemination system.

THREE EXAMPLES OF COOPERATION. J. Leynarie, High Authority of the European Coal and Steel Community (E. C. S. C.), Luxemburg.

I. Medical, Mining and Iron and Steel Documentation Pool. Created on the initiative and with the financial and practical help of the High Authority, it is destined to improve the medical information of research institutes in the European Community specialized in the field of pneumoconioses, dust prevention and suppression of burns. British experts were associated with the work of the pool The research institutes share the tasks of scanning periodicals, wr ting abstracts and translating; they put all information at the disposal of the specialists in their countries. II. Translation-pool of texts in difficult languages, destined to promote the translation by specialized institutes in the Six Countries of technical coal- & steel-literature published in languages of Eastern-European countries and the Far East. For the iron and steel sector, the research institutes have been grouped in a "Association européenne pour l'échange de la littérature technique dans le domaine de la Sidérurgie" (A. S. E. L. T.). An agreement linking the High Authority to the specialized institutes specifies, on the one hand, the terms and conditions for authorizing subsidies to encourage translation work, and on the other, the obligations of those who benefit by it. III. Project for an automatic documentation-pool in the iron and steel sector. In order to improve their exchange of information, the specialized institutes, who are members of A. S. E. L. T. aim, with the aid of the High Authority, to make increased use of electronic machines. A working-group is preparing a thesaurus of key-words in the metallurgical and iron and steel sectors, taking into consideration all the terms of the existing classifications as well as the lists of key-words known at the present time. The thesaurus will be established in five languages, which will all be registered on an IBM 360 and linked together, after which the machine will be able to provide, for a given subject in particular, all the references to articles in periodicals analyzed in the various languages.

COMPARATIVE COSTS OF DOCUMENT INDEXING AND BOOK CATALOGING. L. H. Linder, Aeronutronic, Division of Philco Corporation, Ford Road, Newport Beach, California, U.S.A.

Costs of storing and retrieving information are not well understood, and little information is available in the literature to provide this understanding. This study was undertaken in an effort to determine and compare true costs of indexing reports and cataloging books. In a prior study the author reported an average input cost of $2.99 per report indexed into a machine document address storage system. Indexing was to a depth of 12.6 access points per document. A similar study of traditional book cataloging costs in the same organization is reported here. Traditional book cataloging techniques using Library of Congress Cards, the Library of Congress Subject Heading List and standard card catalog cabinets were applied to a collection of 10,800 books and monographs. Labor costs ($36,683), supplies ($2,312) and equipment amortization ($641) totalled $39,624, or an average of $3.67 per item, an amount similar to that encountered for indexing reports. However, since the system contained a significant number of duplicate copies, the cost of cataloging a book for the first time averaged $5.68 while added copies were accomplished for an average cost of $1.42 each. A further difference is noted when access points are compared. For the report indexing system the average cost per access point was less than 2¢ while in the book cataloging one it averaged nearer 9¢. Since the two systems deal with significantly different sized units of knowledge, and because the associated subject authority control established for the book catalog was a more time-consuming task than the comparable vocabulary control used for report indexing, it is concluded that book cataloging in situations as described above is inherently a more expensive procedure than is coordinate indexing. The final evaluation of the two systems would, of course, depend significantly upon retrieval effectiveness and retrieval costs of the two methods.

THE USE OF SECOND ORDER DESCRIPTORS FOR DOCUMENT RETRIEVAL. Miles A. Libbey, The MITRE Corporation, Bedford, Massachusetts, U.S.A.

It is proposed that one way to increase the efficacy of document retrieval is to define to the computer the descriptors used to index the file. A computer program written in COMIT to implement the proposal and to facilitate testing its capabilities is described. Definitions are given to the computer as a string of terms called a "definitor." These terms, which act as "second order descriptors," are not normally those used as file descriptors. Their introduction provides a controllably broader base for link-finding and match-counting operations by the computer. It also makes possible such things as introducing new terminology and biasing existing descriptor indexes towards special interests or languages without having to re-index the file. The program computes a "pseudometric distance" between a query and each document and prints an ordered list of those documents closer to the query than some chosen cut-off value. (Large files would probably require some preselection, such as that which would result from use of a concordance.) It then substitutes for each descriptor its definitor and repeats the above process. The result is that the subjective human judgement required to evaluate the efficacy of introducing the definitors is reduced to a statement as to which list would be considered more useful. Use of the program to date has been only as a demonstration so no conclusions can be stated other than that the demonstration results would seem to indicate that testing on a serious scale should be undertaken. (This paper is a result of work sponsored by the MITRE Corporation's Educational Assistance Program.)

A THREE-DIMENSIONAL ARRAY USEFUL FOR MAPPING IN FOUR-DIMENSIONAL, NON-COMMUTATIVE FUNCTION-SPACES. K. J. Lissant, Petrolite Corporation, 369 Marshall Avenue, St. Louis, Missouri 63119, U. S. A.

In a recent article[1] a detailed description was given of the use of non-commutative composition spaces for the plotting of properties of polymeric materials. Several alternative geometric arrays were illustrated for plotting in spaces of up to three dimensions. Since that time, a method has been developed and used in our laboratories for depicting four-dimensional composition-spaces. The method is general and should be useful for the study of any four-dimensional function-space. The model to be described in this paper, although three-dimensional, has many of the properties of its analogous four-dimensional space, including a place to "put" the fourth dimension. It can be considered a three-dimensional prospective drawing of a four-dimensional space. It has proved very useful in discussions of four-dimensional problems with less mathematically inclined persons.

(1) Lissant, K. J., "A Unified Method of Delineating Polymeric Species," Journal of Chemical Documentation, 3, 103, (1963).

CURRENT FID ACTIVITY AND FUTURE PROSPECTS FOR THE UDC (UNIVERSAL DECIMAL CLASSIFICATION). O.A.Lloyd, FID, 7 Hofweg, The Hague, Netherlands.

International responsibility for maintaining and revising the UDC rests with the FID's secretariat (The Hague), whose small Classification Department handles all routine UDC business and carries out the general program of FID/CCC - Central Classification Committee for the UDC. Although the FID secretariat normally does not itself publish UDC editions, it controls all proposals for updating and revising the UDC, issuing P-notes to subscribers in many countries, with a 4-month period for comment: P-notes, unopposed or after agreed amendment, are authorized by FID and included in the half-yearly cumulative "Extensions & Corrections to the UDC". Specific subject areas are covered by a network of main UDC committees with working groups, such as FID/C 3 Social sciences (Dutch-German secr.), FID/C 55 Earth sciences (U.S.secr) and FID/C 69/72 Building & architecture (IBCC, with joint FID and CIB participation). It is expected that more than 30 such main committees will eventually be required to fulfil the CCC's longterm program, which aims to redevelop unsatisfactory schedules on more modern lines and introduce new 'common categories' in the auxiliary tables. Vacation of the class number 4 (by merging Linguistics with Literature in 8) has opened the way, after the prescribed 5-10 year 'starvation' period, for a planned new class spanning the pure and applied sciences in classes 5 and 6 and perhaps bridging the gap between social sciences 3 and science & technology 5/6. In applying more facet analysis and notation to future revision work, the CCC will seek help from FID/CR - Classification Research committee, while a joint subcommittee of the CCC and FID/MSR - Mechanized Storage and Retrieval - still preoccupied with studies of mechanized production of UDC schedules - will later investigate the feasibility of using, or adapting, the UDC for machine indexing and retrieval.

A CURRENT AWARENESS PROGRAM FOR THE FIELD OF WATER RESOURCES. Bill L. Long, Chief, Earth Sciences Branch, and David F. Hersey, Deputy Director, Science Information Exchange - Smithsonian Institution, 1730 M Street, N. W., Washington, D. C., U. S. A.

A comprehensive program to provide information about ongoing research in the field of water resources has been developed by the Science Information Exchange (S.I.E.) of the Smithsonian Institution. Rapid and unique solutions to the problems of current awareness in the water field are required due to the intense national interest in this critical area, the highly interdisciplinary nature of the subject matter involved, the recent proliferation of research activities, and the fact that a 1 to 3 year gap exists between inception of the research investigations and initial publication of findings. As the newly-established national cataloging center for water resources research, S.I.E. classifies and maintains for general use a catalog of water research activities as well as providing, on request, information related to specific investigators, institutions and agencies, and research topics. In addition, all new research proposals received by the Office of Water Resources Research, Department of the Interior, for funding are referred to S.I.E. which returns related projects and proposals as an aid in proposal evaluation. Each of these tasks has required the development of special documentation techniques. Requirements for the development of a more automated system for cataloging both for publication and non-publication purposes have been studied.

A CLASSED THESAURUS AS AN INTERMEDIARY BETWEEN TEXTUAL, INDEXING AND SEARCHING LANGUAGES. Gertrude London, Rutgers, The State University, Graduate School of Library Service, New Brunswick, N. J., U.S.A.

Many search problems and failures in the information retrieval field are caused by misunderstanding the purpose of "indexing", especially the part played in it by language with its infinitely diversified forms and meanings. The indexer's task is entirely use- i.e. retrieval-oriented: he has to guide inquirers to the information content of documents, by reducing the multiplicity of word-expressions to categories and transforming these into index-symbols. His means are twofold: (1) words and phrases (subject headings, descriptors, etc.),(2) numbers and letters (classification schemes). In the first case difficulties arise when the symbolic function of the index words or phrases is overlooked because of the nearness of the index symbols to textual words. In the second case the notational significance of the numbers or letters is obvious, but their meanings are not always clearly conveyed in the general statements intended to explain them. Interpretation difficulties in indexing and searching have in recent years led to the compilation of thesauri, some of which try to standardize terminology. The recognition of the complexities inherent in language expression should not induce us to demand that every author use the same words for similar topics, nor that all indexes be confined to the same symbols. We might instead try to create instruments that would help to convert textual or inquiry languages into index-symbols (of various kinds) and vice versa, and that would also act as "translators" between different indexing languages. This translation function might be carried out by a thesaurus, consisting of clearly defined terms, grouped in systematically classed and alphabetically sequential tables, similar to Roget's Thesaurus. The structure of this thesaurus would incorporate two different, but complementary search methods: the "topic-word" and the "synoptic survey or topic-in-content"approach, whose combination might help to deal effectively with different user situations.

LA CLASSIFICATION DE DEWEY et LA CLASSIFICATION DECIMALE UNIVERSELLE. Georges F. LORPHEVRE, Centre National de Bibliographie. Membre du C.C.C. (FID), Bruxelles, Belgique.

C'est en 1895 que MM. Paul OTLET et Henry LA FONTAINE, fondateurs de l'Institut International de Bibliographie (devenu Fédération Internationale de Documentation) adoptèrent la Classification Décimale. A l'occasion de la première conférence internationale de Bibliographie (septembre 1895), ils présentèrent un fichier bibliographique entièrement classé suivant la Dewey. Cette expérience, portant sur 400.000 fiches leur prouva que des adaptations étaient nécessaires, ils proposèrent à Melvil Dewey un certain nombre de perfectionnements, portant, en particulier, sur la création de tables de divisions complémentaires permettant d'obtenir une grande flexibilité dans le classement. Par le fait même la toute première classification à facettes fut inventée. Avec l'accord de l'inventeur, ils publièrent tout l'abord des extraits des tables, puis des adaptations pour en arriver enfin à une version européenne dans laquelle on retrouve, certes, la base commune mais considérablement développée pour servir à des fins bibliographiques.

La collaboration entre Dewey et les fondateurs de la Fédération Internationale de Documentation fut très étroite jusqu' à la première guerre mondiale. Elle fut reprise directement après la fin des hostilités. A ce moment Dokas Fellow y fut intimement liée. Puis les circonstances firent que les besoins de plus en plus précis de la documentation obligèrent les européens à développer les tables originales. En fait deux versions différentes naquirent.

THE DEVELOPMENT OF A MACHINE-SEARCHABLE INDEX-ABSTRACT AND ITS APPLICATION TO BIOMEDICAL LITERATURE. Lois Lunin, Drexel Institute of Technology, Philadelphia 4, Pa.

The purpose of the study was to develop a format for an index-abstract that could be searched by computer and would provide data in the form of key facts contained in biomedical documents as well as references. Pathologists determined those aspects of the literature for which they wished information recorded. A logical sequence of terms for these aspects was developed which resulted in the creation of a miniature abstract (mini-abstract). A thesaurus was prepared for the standardization of terminology in these mini-abstracts. 200 documents representative of the literature on trace metals as related to pathology were indexed by this technique. A computer program was used to store the information on magnetic tape, to search, and to print the retrieved mini-abstracts. Results include 1) the production of a pilot edition of the Trace Metal Literature Index-Handbook which contains groups of mini-abstracts and reference numbers arranged under 90 subject headings, bibliographic citations, and alphabetic list of all authors, and 2) information about the uses of the mini-abstract technique. The mini-abstract meets most of the criteria established for indexes and abstracts; the majority of the papers in the sample require from 1 to 5 mini-abstracts for complete content coverage; the time and effort involved in preparing a mini-abstract depends on the clarity of the original paper. The method can be taught easily. The technique appears most successful with papers which show a cause-effect pattern or those where a comparative statement can be made. The combination of the format of the abstract and the capabilities of the computer permit the dynamic handling of material providing for machine storage and search, for data retrieval in addition to reference response, and for a permanent, easily updated handbook which can be used by researchers at their desks or in their laboratories.

PRACTICAL COMPUTER PREPARATION OF BOOK INDEXES. Clifford J. Maloney and Marvin Katz, National Institutes of Health, Bethesda, Maryland, U. S. A.

The preparation of an index to appear in the back of a published book (an internal index) is to be sharply distinguished from the preparation of an index affording entry to bibliographical items (possibly) unknown at the time of search (an external index). The development of an efficient method for internal index preparation by computer could have considerable interest in general but seems particularly advisable under three definite conditions. The employment of computers in the publication of newspapers, books, and other material is now performed on a routine basis. (Proceedings of the International Computer Typesetting Conference, The Institute of Printing Limited, July 1964) Any additional service obtained as a by-product would, of course, add to the attractiveness of computer typesetting. The publication of reports and congresses such as this one, in which earliest possible appearance of the published proceedings adds greatly to its value, but where author indexing is not practical or, because of the large number of authors of varying backgrounds, is, at the least, very difficult, has more to gain than would be the case in more conventional book publishing. A third field of usefulness exists in rendering indexing possible for material not now indexed. While a few major newspapers have in the past been able to provide indexes to their contents, it is not usual to prepare them and it is not likely to become so until some method can be found to make indexing a by-product (not necessarily entirely automatic) of original publication. The features which distinguish internal from external indexing and which make it easier to achieve have been examined in a previous study. (Indexing Reports by Computer, Clifford J. Maloney, James Dukes, and Sterling Green, TECHNICAL PRECONDITIONS FOR RETRIEVAL CENTER OPERATION edited by Benjamin Cheydleur) One major difference lies in the desirability of very detailed indexing, yet little space can be allowed to print the results. Study of this problem has led to the development of what we believe to be a practical scheme, which has been tried on sample material. We are now seeking to apply it to an actual book.

TEXTUAL INFORMATION RETRIEVAL AND THE IBM SUGGESTION PLAN Joseph J. Magnino Jr, Thomas J. Watson Research Center, CHQ ITIRC, P. O. Box 218, Yorktown Heights, N. Y U.S A

The IBM Technical Information Retrieval Center provides a computerized in-house information retrieval and dissemination system for the scientific and technological community of IBM It is in operation using a textual information retrieval system for retrospective searches, current awareness, monthly bulletins, and special library tools, of IBM technical documents and selected non-IBM documents. Over 120,000 technical documents are on magnetic tape for this centralized service A specialized application of the normal text information retrieval system is using this system to search IBM Suggestions, to provide the necessary capability to determine whether or not a particular suggestion is new or is one that was previously submitted. The system, tested on searches of 40,000 suggestions, uses computer and teleprocessing techniques to assure that 200 searches (new suggestions) are accomplished daily within a 24-hour turn-around time A description of the normal text I.R. search system and the operational procedure, using the IBM Suggestion plan as a special application, is detailed in this report.

ARRANGEMENT OF PRIMARY PUBLICATIONS. D. J. Maltha, Director of the Center for Agricultural Publications and Documentation, Wageningen, Holland.

Primary publications constitute the basis of documentation of the sciences. The greater the uniformity of these publications in the arrangement and development of topics, the easier is the task of recording documentary data from them. Various attempts have been made to increase this uniformity, but these efforts have met with resistance from the authors because they feel that their personal freedom to publish is hampered by the observance of rigid rules. It would, in fact, reduce the level of scientific output if the author's ideas were to be confined in a straitjacket. But it is a generally accepted requirement that in a primary publication the research worker should follow the shortest possible path between the exposition of his problem and the conclusions, and that he should pursue a logical sequence of thought. This means that the same arrangement can be recognised in many primary publications and that a number of fixed points may be indicated in this arrangement. It would be extremely useful for the processing of documentary data from primary publications if these points could be readily identified. It is therefore proposed to mark these points by means of symbols in front of the lines or in the margin. The requirements that such symbols should comply with are being discussed.

CHEMICAL SUBSTRUCTURE SEARCHING WITH LINEAR NOTATIONS.
Beatrice A. Marron, Gloria R. Scheetsky, and Stephen J. Tauber, National Bureau of Standards, Washington, D C 20234, U S A

Among the requirements for a comprehensive chemical information storage and retrieval system is the ability to search among a file of chemical structures both for individual compounds and for classes of compounds defined by any desired substructure. Chemical structures are shown in written communications as networks of atoms and bonds, but linear notations consisting of sequences of symbols are preferred for computer input. A routine has been developed, for the PILOT research computer facility, using Hayward notations (H.W. Hayward, U S Patent Office Report no. 217) to find either individual complete compounds or simple but common types of chemical substructures. The technique relies on matching a symbol string characteristic of the substructure sought (or an entire notation) with strings within the complete structure notations. Provision exists for interruption of the symbol string by arbitrary substrings and for a special type of permutation associated with cyclic substructures. Hayward notations for structures are entered in ASCII (American Standard Code for Information Interchange), together with serial numbers and control codes that aid in limiting errors, and are packed for storage. Questions are entered similarly, either at the time of search or into a file to be called from by serial number. For a search the question is unpacked into separate characters. Each file notation in turn is unpacked, and the abbreviations used in the Hayward notations are expanded. A question may contain any number of alternative acceptable strings of characters, thus providing the Boolean "or". Points within the strings where arbitrary substrings are acceptable are labelled by a "don't care" character. Permutations of strings are made on a substring by substring basis. This search routine is intended to be one module of a comprehensive chemical information handling system. It provides one route to structures, the most common access point to chemical information.

[redacted]

ON THE ECONOMICS OF COMPUTER STORAGE AND RETRIEVAL
H. Marron, Assistant Director, and M. Snyderman, Deputy Assistant Director, Science Information Exchange - Smithsonian Institution, 1730 M Street, N. W., Washington, D. C., U. S. A.

Despite the widespread interest in the economics of computer storage and retrieval activities (specifically the cost of performing searches and preparing reports) there is little definitive discussion of this subject in the open literature. A few operating installations have released data on unit search costs obtained by dividing total computer processing time for a batch of jobs by the number of search questions in the batch. By so doing, however, the costs for separate single question searches or for building, modifying, maintaining and updating the file are neglected. These cost may be large and can hardly be disregarded!

Alternative techniques have been studied for costing unit searches which include these factors: (1) Computer times for all searches batched or single (2) Proration of the cost of a batched search to each of its component parts and (3) All computer times for file modification or maintenance tasks. Theoretical results have been compared with actual operating data from computer searches performed at the Science Information Exchange. Further, some generalized relationships have been developed which can be used by managers, analysts and designers for qualitative and quantitative assessments of the impact of file modification and maintenance task on search costs.

This paper confines its attention to direct computer costs only. Later papers will deal with the overall costs - which include programming, analysis and overhead.

"WHERE CAN WE GET COMPETENT HELP?"
John Martinson. Inst.for Advancement of Medical Communication 9650 Wisc. Ave. Bethesda, Maryland. U.S.A.

"All chiefs and no indians" is a common complaint. Yet, the clear recognition and demarcation of professional from technical or semi-professional tasks in the information sciences is only beginning to emerge. If professionalization in this field follows the pattern of others, sharper delineation of responsibilities will occur. (Many analogous examples can be drawn from medical fields.) -- The growing demand for library personnel has brought such terms as LIBRARY TECHNICIAN or LIBRARY TECHNICAL ASSISTANT into wider use. This development has been formalized in the past decade (and largely within the last 5 years) through establishment of special curricula to prepare students for jobs as Library Technicians. Approximately 25 schools in the U.S. offer such courses. These are mostly 2yea programs leading to an Associate of Arts degree from a Community or Junior College. -- Through site visits, phone interviews and correspondence these training programs have been surveyed to answer a number of questions. These include:

who is recruited into these programs and by what means?
what training and experience do they receive?
what is the basis for the design of the curriculum?
what are the chief problems in the administration of these programs?
where have Library Technicians been placed and what is the nature of their employment experience?
what future problems are anticipated by those responsible for the development of these programs?
what lines of research are indicated to help solve present and anticipated problems?

These programs are often characterized by all the difficulties of rapid growth in an effort to meet explosive problems with rather meager resources. The growth rate of these programs appears to be increasing, however, much work on curriculum development needs to be done. Many examples of successful employment can be cited. -- Copies of the Final Report to the U. Office of Education will be available. Oral presentation will focus on discussion of selected examples of Recruitment; Training; and Employment experiences.

THE COVERAGE, INDEXING AND OVERLAP OF ABSTRACTS SERVICES. John Martyn and Margaret Slater. Aslib, 3 Belgrave Square, London SW 1, England.

Studies of the coverage and subject indexing of abstracting and indexing services have been made using the technique reported in J.Doc 20 4 (Dec. 1964) pp. 212 - 235. Comprehensive bibliographies on specific subjects are taken, and the appropriate abstracts journals examined to determine coverage, and overlap between abstracts journals, of the references in each bibliography. Then the subject indexes of the abstracts journals are examined to find the terms under which the covered references were indexed. Figures are presented for the coverage and indexing of the following subjects : Magnetohydrodynamics, Evaporation control, Determination of tritium, Radioactive fallout monitoring, Motivation research, The red cell, Magnetic and electric suspensions and Biological control. Abstracts services covered include The ical bstracts, Science Abstracts (A & B), Biological Abstracts, Index Medicus, Engineering Index, Nuclear Science Abstracts, and others.

A PROPOSAL FOR REDUCING THE TIME LAG IN SCIENTIFIC PAPER PUBLICATION SCHEDULES. Hunter P. McCartney, IBM Federal Systems Division, 7220 Wisconsin Avenue, Bethesda, Maryland, U.S.A.

The lag between receiving and publishing papers by scientific journals acts at cross-purposes with these journals' objectives. Authors hesitate to risk being "scooped" during this long period and resist the obsolecence a year can add to their subjects on the other (a year is about the average time lag under our current schedules). Those concerned with a paper's publication - the author, the reviewers, the editor, the publisher - are well aware of this excessive lag, but not much is being done about it. The seriousness of the problem is shown by the proliferation of letter-type publications (Applied Physics Letters is a prime example) established to assure authors earlier exposure of timely subjects. But the advantage of faster publication by these "letters" and by the "correspondence" sections of journals could be offset by a lower threshold of quality, for to meet these accelerated publication schedules, editors are forced to settle for less than the preferred rigorous review they arrange for normal papers. Much of the publication lag stems from a paper's review, for most reviewers sandwich this time-consuming, and often onerous, work into crowded schedules and frequently delay these gratuitous efforts because of their job demands. Another delaying factor is the shortage of the specialized talents and facilities needed to set articles in type. This paper is built on the author's observations, on conversation with editors, authors, reviewers, and others concerned with the timely dissemination of scientific information, and on selected articles. No simple remedy is proposed, but improved reviewer reward, an increase in special facilities and reproduction talents, and more attention to origination and presentation, as well as to dissemination, would provide short-term relief. The ultimate solution to these problems is to depart from our present methods of publication and to use the computer more efficiently in publishing, classifying, and disseminating information.

PRECOORDINATION AND SELECTIVE PERMUTATION. Nan P. McCandless, Pacific Aerospace Library, 7660 Beverly Blvd., Los Angeles 36, California.

The precoordination of multiple associations printed out in order to provide immediate access to the entire store of an information system will provide a high degree of specificity. This may increase the volume of the printout and also, when the permutation printout requires several pages for a single term, make it difficult to relate the associated descriptors together. There are various methods to correct this situation, one of which is selective permutation. By this method all descriptors which are linked together by coincidence of the accession number are no longer permuted indiscriminately, and various criteria for selection can be adopted.

1. Selective permutation based on type of descriptors or descriptor ranking. In this case two or several descriptor variations are adopted, they may be called major or minor descriptors; if more than two categories are considered, a numeric scale of descriptor values may be adopted. By this method it will be possible to select the type of descriptor which will be permuted. It will also be possible to permute as major look-up descriptor or index entry any of these types of descriptors, or include in the associative permutation any of these categories, without having them as major index entries.

2. Selective permutation may also be based on a combination of descriptor and classification or group identification. By this method the alphabetical sequence of the associated descriptors may be subordinated to one or several major groupings. By this method it is more likely that descriptors with present similarity in concept or function will appear in the same printout area.

The precoordination of the Pacific Aerospace Index provides statistical observations for the alphabetical precoordination, and also for selected areas, practical considerations for the precoordination and permutation of the two alternate methods considered above.

TEXT90. V.S.Mercer and F.E. Franklin, Automated Documentation, Department D78, Systems Development Division, International Business Machines Corporation, Poughkeepsie, New York.

TEXT90 represents a major breakthrough in automated text preparation. An IBM 7090/94 program, it accepts free-form text data and formats this data into "camera ready" copy that is suitable for reproduction. Through easy-to-use controls, the user can specify formatting of a line, page or the entire document. TEXT90 has both update and format capabilities. The updating capabilities range from the deletion, insertion or replacement of a character, word or line(s), to the moving of entire blocks of text from one area of the document to another. This gives the user complete control of the document during its various stages of preparation. Formatting capabilities provide for normal functions (e.g., skipping, indenting) and complex functions (e.g., table and figure generation). These capabilities, when combined with double or single column format, automatic hyphenation and justification, assure the user of a properly formatted document. TEXT90's freeform input can be prepared on punched cards, or the data stored in the Administrative Terminal System can be converted to TEXT90 using the ATS-TEXT90 conversion program. TEXT90 output is printed on an IBM 1401-1403 equipped with a 120-character upper/lower case print chain. TEXT90 has been proven. Thousands of pages...hundreds of thousands of lines...have been published using TEXT90. It has made possible the rapid dissemination of up-to-date documentation so vital to the success of the IBM System/360 programming effort.

SYSTEM ANALYSIS OF SCIENTIFIC INFORMATION IN THE FIELD OF BASIC RESEARCH. Augustin Mertа, Czechoslovak Academy of Sciences, U půjčovny 5, Prague, Czechoslovakia.

Object of research: theoretical, methodological & organizational problems of SI (scientific information) in the institutes of natural, technical & social sciences of the Czechoslovak academy of sciences. The purpose of research: creation of effective, coordinated network of SI centers. Methods used: statistical questionaries, discussions with SI workers, managers & senior staff scientists. Attention to correlation of the personal & institutional documentation systems. Results obtained: effectivity of managing, rationalization & economization of scientific work is in direct proportion with informational level of managers & scientists. Heuristic & methodologic differences amongst scientific disciplines do not allow to create unique all-science SI system. Factors limitting individual SI projects: extent of problems to be solved, ratio of experimental, theoretic & information work, necessity & possibility of using external sources of SI, application of scientific results in the sphere of technology, existence & attainability of relevent abstract journals, indices & other secondary resources, useful life of actual SI. Scientists cannot be excluded from the information process. He should indicate what information has to be put into memory of SI system & cooperate in creation of effective indexing plan. It is necessary to raise theoretic & technical level of SI work, aswell as qualification of information scientists & other SI workers. The problem of centralization or decentralization of SI workplaces needs more investigation. Having introduced abstracting & indexing machines synchronized with effective telecommunication means, it seems possible to build up central SI institutes.

FUTURE ROLE OF THE REACTIVE TYPEWRITER IN DOCUMENTATION. Calvin N. Mooers, Rockford Research Institute Incorporated, 140½ Mount Auburn Street, Cambridge, Mass. 02138, U.S.A.

Within three years the reactive typewriter will be coming into use in documentation in USA and overseas. The reactive typewriter is the successor to the tape typewriter, with a telegraph line in place of the paper tape. Specifically, the reactive typewriter is a teletypewriter which acquires powerful new capabilities through on-line connection to a nearby multi-access time-shared computer having large-scale storage capabilities (C.N.Mooers, "The Reactive Typewriter Program", Comm ACM v.6 p.48 Jan.1963). Between twenty and several hundred reactive typewriters will be connected to the same computer, sharing the cost. The cost is expected to be less than $300 per month, complete. They will be widely used, replacing many ordinary typewriters in laboratories and libraries. We estimate that at least 500 are now in use. Applications in documentation are beginning. Typical uses include manuscript preparation, business letters, library cataloging, union list preparation, transmission of bibliographic information, as well as computation. Users of the reactive typewriter will set up private indexing and retrieval systems, with retrieval being performed from their typewriter. Using communication facilities (e.g., TWX and Telex) similar access will be had to various documentation files in national and international bibliographic networks (cf. de Grolier in "L'Organization de la Documentation Scientifique" ed. Poindron, pp.93-154, Gauthier-Villars, Paris, 1964). We are developing a machine-independent computer programming language called TRAC for controlling the interaction between the user and his reactive typewriter. TRAC has been running under test for 1½ years. With TRAC, from the reactive typewriter keyboard, one can (1) store, make any changes in, and type out any text; (2) store and index collections of entries; (3) store procedures for doing things; and (4) access external communication channels. (cf. C.N.Mooers, "TRAC, A Text Handling Language", Conference Proceedings, ACM National Conference August 1965, to be published.)

INFLUENCE OF THE NEW GENERATION OF COMPUTER SYSTEMS ON LIBRARY AUTOMATION AND INFORMATION RETRIEVAL. Thomas L. Minder, International Business Machines Corporation, Federal Systems Division, 7220 Wisconsin Avenue, Bethesda, Maryland, 20014, U.S.A.

Changes and improvements in computer hardware in recent years have caused changes in the design of information systems. This paper is a discussion of the affects of increased computer core capability, direct access, shared time, remote stations, modular design and larger printing font capabilities on the design of library and information systems. The discussion will highlight the effects of these properties on classification systems, filing and sorting rules, throughput time, computer querying vs output listings, standard software packages and unit record concepts. The changes in personnel qualifications will also be noted.

DISSEMINATION OF JAPANESE NUCLEAR LITERATURE IN THE WORLD. Taisuke Nagayama, Japan Atomic Energy Research Institute, Tokai-mura, Naka-gun, Ibaraki-ken, Japan.

The dissemination of Japanese nuclear literature involves many problems, because of the language barrier and the way of its dissemination. An investigation has been made of the subject matter, using the papers whose abstracts have appeared in the Nuclear Science Abstracts of Japan. Approximately 2000 papers, covered in the above abstract journal in the past two years, were analyzed; as a result, the Bradford's chart was prepared. Thirty journals were selected as the core and circumferential ones, and their dissemination in the world has been analyzed in detail. Forty percent of all the papers have been written in English; of these, 60% were carried in English-language journals and 40% in Japanese-language journals. The percentage of papers written in English has differed from one field of research to another, 70% for physics, 30% for chemistry, etc. Abstracts in English were attached to 80% of the papers. In the case of English-language journals, there is a problem of correcting the English used in these papers. These journals are intended for overseas distribution, and 20 to 30% of the numbers of copies are sent to overseas countries. The distribution differs considerably from one country to another, with the United States receiving nearly 50% of the copies sent. For Japanese-language journals, on the other hand, only 1 to 2% are sent. Reprints are playing a role in the dissemination of papers. Thus, the state of dissemination of papers in the world is considered unsatisfactory. International abstract and index journals are made use of in this connection. Papers written in English are then easily accessible. For the papers written in Japanese, however, there is a problem of translation.

ECONOMY IN THE IDEA PLANE IN THE DESIGN OF A SCHEME FOR DEPTH CLASSIFICATION. A Neelameghan and M A Gopinath, Documentation Research and Training Centre, Bangalore 3, India.

Three concurrent developments in Indian classificatory thought are: Investigation on the methodology of design, making the design work amenable to scientific method, and making the work of the classificationist and of the classifier more productive. Productivity in design is increased by the separation of the work in the idea, verbal, and notational planes, bypassing the confusing phenomenal level by diving deeper to the near-seminal level to sense the essence of the bonds, assigning a scale of priority to the problems to be dealt with, formulating guiding principles, and reducing the strain on the memory by the use of mnemonics. Discusses the residual problems in the application of the Wall-Picture Principle in determining a helpful sequence among the First Characteristics in the schedule for the Personality Facet in subjects going with a Basic Class in Commodity Production Engineering. Observes a pattern in the formation of groups with internal affinity among the First Characteristics. In step 2 of design suggests such grouping of the First Characteristics, the application of the Wall-Picture Principle first to the groups to determine a helpful sequence among them, and then to the First Characteristics within each group to determine a helpful sequence among them. This nearly eliminates the flair necessary to select pairs of First Characteristics one after another for the direct application of the Wall-Picture Principle to one pair at a time, and also reduces the time taken to determine a helpful sequence among the totality of the First Characteristics. Further, the groups of First Characteristics with internal affinity are recognised even by a beginner without much difficulty. The groups of First Characteristics with internal affinity lie at a level nearer to the seminal level than that of those of the First Characteristics themselves. The Wall-Picture Principle gives a sequence of the First Characteristics which nearly parallels the sequence of thinking by the majority of the production engineers.

STATUS AND CONTROL OF UDC FULL EDITIONS. D. Newcombe, British Standards Institution, 2 Park Street, London, W.1., England.

A survey is made of existing UDC full editions, i.e., in French, German, English, Japanese, Spanish, Portuguese and Polish, and of projected full editions, e.g. in Russian, and a tabulated comparison of the classes, divisions and sections available in different languages is provided. The international and national organs for the control, revision and extension of the UDC are described giving an indication of the purpose, structure, membership and activities of such organs, i.e. the Central Classification Committee and its special sub-committees and working groups, the various national UDC committees and their sub-committees and working groups. Finally, an outline is given of the existing Rules for UDC Revision Procedure: general principles, classification and notation principles, preparation and submission of proposals, treatment and circulation of draft proposals, publication of proposals as P-notes, conversion of approved P-notes into "Extensions and Corrections to the UDC."

THE HYPERTEXT. Theodor H. Nelson, Vassar College, Poughkeepsie, N.Y., U.S.A.

We are at an historic divide like that created by movable type. To supplement movable type came such further inventions as the footnote, preface, magazine, bookshelf and editor. The technology of automatic display, stored-program computers and bulk information storage call now for a like range of inventions and conventions: display modes, organizing arrangements, and academic and professional roles and practices that will give us the most from these new devices.

This paper proposes a new medium, the hypertext. This is a generic term for texts (and combinations of texts with other materials) which, because of their structure, require automatic handling and display devices. The hypertext will typically be non-linear, branching, and large, with various options to the user.

The idea of the hypertext is distinct from Information Retrieval (specialist dissemination and query search), Information Display (presentation of data desired in a known context) and Programmed Education (restricted-sequence presentation and drill). None of these fits the general case. The hypertext is suited not just to specialty materials (which fit on paper nicely), but to give general reference and instruction. It may contain a corpus of general and special materials, with their many interconnections indexed. Arbitrary boundaries of subject matter can be ignored. Hypertext may be built out of existing writings, not mechanically heaped but personally assembled, according to considerations of scholarship and pedagogy. Like any anthology, the hypertext may contain interrelated texts and information from diverse original sources. It may also contain summaries and abstracts written at different levels, permitting separate entry to users of different backgrounds, competences and interests, allowing them to pursue such details or overview as they desire. A hypertext system could provide instant lookup of definitions, biographies, or explanations. It could have new types of graphic material (moving or anamorphic), ways to show relations (indexes, footnotes, marginalia), even of browsing or reading (e.g., stroboscopic). Discursive, conceptual and general-purpose, it will permit self-education according to interest and motivation, even exploiting natural curiosity and intellectual emotion, which so often hinder conventional instruction.

COOPERATION AT AN INTERNATIONAL LEVEL AS A MEANS OF ORGANIZING DOCUMENTATION NETWORKS. T. F. Noyon, High Authority, European Coal and Steel Community, E. C. S. C., Luxemburg. R. Brey, European Atomic Energy Community, Euratom, Brussels, Belgium.

The constant growth of scientific and technical literature is giving difficulties at all levels: information center, individual research worker, management of research. The use of computers can help, but cooperation and sharing of tasks are also important. This can be organized vertically, within a national framework, with a multipurpose center at the top, or horizontally, at the level of centers with similar specializations in different countries. The latter form of organization promises to hold the key to the problem, but the former must not be ruled out, because a multipurpose center can, for instance, prevent duplication of work. The horizontal form of organization implies international cooperation, legal links between members and financial resources. In the case of the European Community of six countries, these conditions are fulfilled. For instance, E. C. S. C. and Euratom, can play a federating role in the organization of information on coal, steel and nuclear energy. Cooperation has in particular cases been extended to other countries, e. g. the United Kingdom and the United States. Practical illustrations of this way of thinking are given in two papers. One paper deals with the creation by E. C. S. C. of a pool of medical documentation for the mining and steel industries, a translation pool of technical texts relating to coal and steel written in the more difficult languages and a projected automatic documentation pool in the field of steel-making. Another paper describes an automatic information storage and retrieval project set up by Euratom which will soon be operational and which the United States Atomic Energy Commission is collaborating.

CONTRIBUTION TO THE STUDY OF METHODS FOR CATALOGUING MUSEUM COLLECTIONS. Yvonne Oddon, Unesco-Icom Museum Documentation Centre, Unesco House, Place de Fontenoy, Paris (7e), France.

The cataloguing and classification of museum collections have not yet been properly codified. Two methods are examined in this paper. The first one, meant for general or regional museums of medium size, is based on a visual cataloguing and analysing principle with use of especially designed "Synoptic" cards: a standard descriptive card is filled up for each museum object and placed in the pocket of a Synoptic coded card bearing coloured signs corresponding to the main characteristics of the object. It is thus possible to refer rapidly to collections according to their origin, date, mode of acquisition, material or technique, author or species, state of conservation, location in the museum, etc. The second method refers to the "Filmorex" system, codified micro-cards, which are well suited to the electronic selection of museum pieces. They reproduce both the descriptive cards and the photograph of the object, they can be cheaply multiplied in order to allow for several pre-classified catalogues and they can automatically, by photolisting, give photographic enlargements on paper. Such a method is of course mostly suitable for "mass cataloguing" and for use on a national or international level.

POSIBILIDADES DE CREACION DE PAISES EN DESARROLLO EN MATERIA DE DOCUMENTACION - URUGUAY. Raquel Ortiz Aguier y Maria Rosa Capó. Universidad de la República, Montevideo, Uruguay.

Los países en desarrollo sienten la necesidad de organizarse en materia documentaria. La carencia casi total de medios para hacerlo y otras deficiencias, inciden para que esta actividad se posterme. Se hará un estudio de la situación actual del país y de los medios precisos con [illegible], antes de iniciar otra actividad, habrá que dedicarse a la elaboración del catálogo colectivo nacional, en nuestro medio [illegible] y temido. Siendo el catálogo colectivo la base de [illegible] organización de documentos, proyectaremos elementos que permitan llevar a cabo los siguientes proyectos: racionalización de compras; coordinación con los centros bibliográficos regionales; bibliografía nacional; listas de bibliotecas, institutos especializados e investigadores; investigaciones en curso; listas de revistas nacionales; resúmenes analíticos; reproducción de elementos; cursos de adiestramiento para estudiantes e investigadores; índices de revistas, congresos y bibliografías especializadas; bibliografías de bibliografías; bibliografías especializadas; información y referencia; boletines bibliográficos; propaganda y estadística. Se deben hacer planes de realización inmediata y aprovechar los esfuerzos aislados o esporádicos, coordinarlos y establecer un programa de desarrollo documentario. Los resultados en base a la realización de esta etapa, crearán la conciencia nacional de las necesidades documentarias lo cual se logra con hechos y no sólo con planes. Cuando se alcance a demostrar lo realizado y su utilidad, habrá mayores perspectivas de que las autoridades destinen los medios económicos indispensables para aplicar las técnicas modernas de documentación mecanizada y colaborar en planes regionales e internacionales de documentación.

PEER-GROUP JUDGMENTS ON SCIENTIFIC MERIT: EDITORIAL REFEREEING. Richard H. Orr and Jane L. Kassab, Institute for Advancement of Medical Communication, 4040 Locust Street, Philadelphia, Pennsylvania 19104, U.S.A.

In making decisions on awards, honors, promotions, and publication, the scientific community relies on peer-group judgments of a scientist's work, i.e., on the consensus of colleagues qualified to assess his accomplishments. Despite the importance of this method of evaluating scientific merit, neither it nor alternative methods has been examined objectively. Because the practice of using editorial referees to judge manuscripts submitted to journals provides a model of the general method that lends itself to systematic analysis, and because editorial refereeing is considered to be the primary means of quality control in the documentation chain from generator to user, it was selected for the first of a series of studies on evaluation of scientific merit. The aims of this study are: 1) to assess the editorial refereeing process quantitatively, 2) to identify factors that affect the process, 3) to generate hypotheses on how its reliability and efficiency might be improved, and 4) to compare the process with other means of evaluating scientific merit. Editorial records on all manuscripts (almost 5,000) submitted to two biomedical research journals from 1957 through 1961 were the chief data source. These journals used a form for referees' reports that required a numerical rating of scientific merit (1 = "superior", 2 = "good", 3 = "low but acceptable", and 4 = "unacceptable"). In addition, a sample of the manuscripts submitted in 1957 were returned to the referees who originally reviewed them to be re-evaluated in retrospect. The findings to date include: 1) On the critical decision of whether a paper was acceptable for publication, referees agreed in about 70% of all cases where only two opinions were obtained. 2) There was closer agreement on "poorer" papers (ratings 3 and 4) than on "better" ones (ratings 1 and 2). 3) A referee's retrospective judgment about the acceptability of a paper agreed with his original opinion about as often as two different referees agreed. 4) In their retrospective judgments, referees tended to find fewer papers unacceptable.

STANDARDIZATION OF DOCUMENTATION AND ITS INFLUENCE ON EDITING OF CHEMICAL JOURNALS IN JAPAN. Yasuhiro Ota, Central Research Laboratories, Ajinomoto Co., Inc., Suzuki-cho, Kawasaki-shi, Kanagawa-ken, Japan.

The standardization of documentation is carried out mainly by the International Organization for Standardization Technical Committee 46 (ISO/TC 46), which approved so far tens of ISO recommendations on documentation. Unfortunately Japan has no national standards of documentation. Though the editors of scientific journals in Japan had their own rules of editing according the modern society's needs, they were never conscious of the ISO recommendations. Eighteen leading chemical journals in Japan were examined for the possibility of accepting the ISO recommendations: ISO R215, "Presentation on Contribution to Periodical," is generally received except recording classification marks; the use of ISO R214, "Abstracts and Synopses," is not sufficient, e.g. most of the journals come to introduce gradually synopses in each articles, but resist to have abstract columns; for ISO R77, "Bibliographical References," the bibliographical descriptions of books or other separately published works as references are not standardized, and if contributions to collective works or symposia are used, their descriptions are desperate; ISO R8, "Layout of Periodicals," is generally received, but in some journals there are no existence of title pages, incomplete bibliographical marks on each page, or doubling of pagenations (paginations both for issue and for volume are recorded on each page); ISO R18, "Short Contents List of Periodicals or Other Documents," is received in many journals; ISO R30, "Bibliographical Strip," is scarecely applicable in Japan's journals; and ISO R4, "International Code for Abbreviation of Titles of Periodicals," is never applicable in chemical journals in Japan because Chemical Abstracts is very familiar for Japanese chemists to result to use Chemical Abstracts' code for abbreviation.

TESTING THE EFFECTIVENESS OF A THESAURUS-CONTROLLED SUBJECT INDEX. Paul E. Palatt, American Society for Metals, Metals Park, Ohio, 44073, U. S. A.

The Review of Metal Literature, a metallurgical abstract journal published by the American Society for Metals, has instituted this year a thesaurus-controlled Subject Index. The objective of this study is to determine the content and format of the index so that the user will be provided with an effective retrieval tool. The measure of effectiveness used is the computer output from the 'telegraphic' abstract system employed by ASM over the past eight years. The same file of documents (approximately 3000) was both hand searched through the subject index and computer searched. The search questions were selected from those submitted by subscribers to the ASM Information Retrieval Service and represent a cross-section of all the types of questions received. The study analyzes the reasons a document was retrieved by the computer search of the deep index and not retrieved by the subject index search. Quantitative results concerning the effect of indexing depth on retrieval efficiency are presented. The results of this study have led to changes in subject indexing practice and these changes coupled with subsequent testing are expected to lead to a truly effective subject index.

CATEGORIES AND RELATORS: A NEW SCHEMA. Jean M. Perreault, Florida Atlantic University, Boca Raton, Fla., U.S.A.

In a mechanized system of bibliographical storage and retrieval, the need for a more adequate system of relational terms with which to eliminate the chance of false coordination (and to make possible new complex classifications) between extant indexing/classing terms (whether verbal or notational) led originally to a comparative tabulation of various available relational and categorical schemata. Their omissions seem traceable to their pragmatic (non-systematic) originations; but even when taken as one collated system some of these omissions remain. (Several available schemata are shown and discussed.)

An alternative route to an adequate and universal schema is described: Just as there are signs of interplay between 'general categories' in the setting up of the extant schemata, such an interplay, systematically carried through, is made the basis for a new non-pragmatic schema. The general categories used are a triad of triads: CANONIC (e.g. lateral, axial, vertical); POSITIVE--INDETERMINATE--NEGATIVE (e.g. active, interactive, passive); and TOTALITY--TOTALITY/ELEMENTS--ELEMENTS (e.g. subsumptive, determinative, ordinal). A notation of letters (a--i) or of numbers (1--9) embodies the concepts in a hierarchical arrangement, offering relations from the most general to the most particular.

(A chart abstracting the conceptual types from the various available lists is given as introduction to (1) a fully systematic schema, (2) a classified-index tabulation, and (3) an alphabetical chain-indexed tabulation.)

Examples are given of use of the schema in combination with U D C numbers in indexing narrowly- or broadly-focussed documents, and projections of results of the combined system when used as a strategization for mechanized searching are given in discussion.

AN INFORMATIVE ABSTRACTING TECHNIQUE: DEVELOPMENT AND VERIFICATION. Dan Payne, John Hale, and Sara Munger, American Institutes for Research, 410 Amberson Avenue, Pittsburgh, Pennsylvania, U. S. A.

The goal of the research was to develop and verify guidelines for preparing highly informative abstracts. In terms of objectives, the study was intended to develop guidelines that result in abstracts which (1) provide maximal support to abstract-users, and (2) result in consistent, or reliable, abstracts of scientific/technical material. An Abstract Form and Instructions were developed by integrating features from 132 different sets of abstracting instructions received from requests for such instructions sent to 276 organizations. The form consists of a number of sections and sub-sections, generally related to the organization of journal articles, with specific instructions for each sub-section. Two studies were conducted to verify the utility of the abstracts produced by the guidelines. In the first study, a test of consistency of the abstracts was conducted. This test was based upon expert judgment of similarity of the amount of information contained in each of the sub-sections of an abstract. Six papers were abstracted by three different abstractors. The 18 abstracts were judged to be 88 per cent consistent in terms of information content. Tests of performance support involved comparing performance using abstracts with performance using full text, under two different conditions of time, on tasks known to be dependent upon textual material. The first study, employing 20 senior electrical engineering students, indicated that there was virtually no difference in quality of performance, but that significantly less time was required by students using abstracts. The second study involved four different general tasks encountered by scientists in industry, and included two types of abstracts: the general-purpose, or full, abstract, and an abstract "tailored" specifically to each of the tasks. Subjects were 228 students and 40 professional scientists. Results indicated that both abstracts served about as well as full text, but with significant decreases in time required over the two abstract types.

ORGANIZATION OF NUCLEAR SCIENCE CONFERENCE LITERATURE. Margaret L. Pflueger, Division of Technical Information, U. S. Atomic Energy Commission, Oak Ridge, Tennessee, U. S. A.

Scientific information presented at conferences is becoming a permanent part of the recorded scientific literature. A study of review articles on nuclear subjects shows frequent references to conference papers with no indication of publication. While many conference papers are eventually published in journals, they are often quoted and referenced as papers before they become part of the journal literature. Many conference papers are never published in more permanent form, and the author is the only source of availability. The U. S. Atomic Energy Commission has attempted to cope with this type of nuclear science literature in a number of ways: by acquiring and announcing it in Nuclear Science Abstracts, by making it publicly available, and by using special indexing techniques to aid in its easy identification and location. A unique KWIC index provides access to information on past conferences by date, city, and important words in the title. A discreet numbering system brings together preprints of papers presented at a conference and links them to the published proceedings.

ILATERAL AND MULTILATERAL INTERNATIONAL COOPERATION IN SCIENTIFIC AND TECHNICAL INFORMATION. Wojciech Piróg, Centralny Instytut Informacji Naukowo-Technicznej i Ekonomicznej, Warszawa, Al. Niepodległości 188, Polska.

A concise outline of current trends in, and development of, international cooperation, bilateral and multilateral, in the field of scientific and technical information, with examples cited from several countries and regions. The cooperation of international organizations concerned with problems of scientific and technical information. The necessity of further development of the cooperation in the field mentioned, in order to match the violent progress in science and technology. Mutual services rendered by the national information centres, the character of these services, the exchange of the know-how, the enterprizes organized in common, the assignement and distribution of tasks in documentation. Optimal conditions of cooperation. The basic concepts concerning the principles and forms of bilateral and multilateral cooperation. Bilateral and multilateral agreements; jointly organized and sponsored international agencies and organs. The problem of the international service of scientific and technical information. The role of FID in initiating and coordinating the international cooperation in the field of information and the tasks assigned for the coming years.

ANALYSIS OF CITATIONS IN SOURCE PAPERS IN THREE RETROSPECTIVE BIBLIOGRAPHIES ON (1) HYDROFOIL CRAFT, (2) CONTINENTAL DRIFT, AND (3) TSUNAMIS(SEISMIC SEA WAVES). Sharlene Rafter, U.S. Coast & Geodetic Survey, Library Br., 11800 Old Georgetown Rd., Rockville, Md., 20852, U.S.A.

Analysis of citation patterns in source papers of three retrospective bibliographies, each covering a specific field, with controlled time span, 1900-1964, was made to establish use factors and citation patterns in scientific and technical literature to guide acquisition, retention, and indexing policies in technical libraries. Hydrofoil Craft, Design and Development, 1900-1964, Continental Drift as a theory, 1900-1964, compiled by the author, and Annotated Bibliography on Tsunamis, compiled by U.S. Coast & Geodetic Survey (IUGG Monograph 27) were used for the study. Source cards were made for each citation in the bibliographies. References cited in source papers were checked. Items contained in bibliographies were noted "cited in..." on source card. Items cited, not in bibliographies, were added with reference as first citation. ZATCR edge-notched cards were used as source cards. Analysis of citations shows that: (1) Principle factors in subsequent citation of papers are professional standing of author and journal of original publication, and language of original. (2) "Classic papers" are cited repeatedly, while major portion of published literature is occasionally cited. (3) Papers are cited more frequently during first five years after publication, less frequently during next five years, and sporadically after first ten years. Most frequently cited papers are included in bibliographies, specialized monographs, and reviews of the literature, which are then cited, rather than original papers. (4) Major portion of source papers is originally published in a comparatively few leading journals. (5) Citations are the keys to retrospective bibliography preparation, serving as multiple access points to specific areas of literature. Importance of these five factors to technical library acquisition, retention, and documentation policies is discussed.

SCIENTIFIC AND TECHNICAL INFORMATION FILE FOR THE NATIONAL CANCER INSTITUTE. A. W. Pratt and W. C. White, National Institutes of Health, Bethesda, Md.

A computer-based information processing system is being evaluated as an aid in the management and analysis of the Grants Program of the National Cancer Institute, NIH. The individual record which forms the data base describes one grant and contains administrative, fiscal and scientific data. Each record is composed of an ordered set of nineteen data-line entries; five entries are of fixed-length format, the remainder have a variable-length format allowing a multiplicity of data items to be entered per line. The scientific data are subject-descriptive key terms or phrases abstracted from the grant document by human indexers. Management activities require identification and retrieval of any subset of records relative to any specified logical combination of administrative, fiscal and scientific data, the analysis activities require in addition, the potential for the exhaustive organization by scientific subject content of any identified subset of the file records. A paper tape-oriented, IBM 1620 with 40K digits of core and random access capability plus a printer comprise the machine system. Computer programs provide for: 1) the creation and maintenance of the machine-stored, natural language information file, 2) the retrieval of records by search on specific data items using combined logic operations, i.e. conjunction (AND, AND NOT) and disjunction (OR), 3) the systematic organization of the set or any subset of the file records by an in-depth matching of a record's subject content against a machine-ordered list of subject content automatically abstracted from the entire record set, 4) the routine sorting, alphabetizing of data lists and a variable output format capability

HOW BIOMEDICAL INVESTIGATORS USE LIBRARY BOOKS. L. Miles Raisig, Meredith Smith, Reneta Cuff, Frederick G. Kilgour, Yale Medical Library, 333 Cedar Street, New Haven, Connecticut, U. S. A.

Only a few studies have been concerned with the use of biomedical books. This paper reports an investigation into use made of library books by biomedical investigators. Cancelled charge slips were collected at the Yale Medical Library circulation desk each day. On the following day those slips for books which had been returned by research investigators were segregated. Next, an appointment was made with an investigator for an interview. The interviewer obtained answers from the investigator to questions on a questionnaire designed to elicit information as to how the investigator had learned of the existence of the book, whether or not the book had been useful to him, and if it had, to what use had he put it. During the six-month period of the study, researchers returned 2,735 volumes of which 831 or 30.4 per cent were monographs. Nearly four-fifths of the books withdrawn supplied information wanted and about four-fifths of the books used were printed in the previous decade. Nine-tenths of the use of books by 130 academic investigators was research-related; the other one-tenth is used for lecture preparation. Over a quarter of book usage is to obtain general information and about 15 per cent is associated with the intellectual aspects of scientific activity.

EDUCATION FOR DOCUMENTALISTS. S R Ranganathan National Research Professor in Library Science Documentation Research and Training Centre Bangalore 3 India.

The umbral region of documentation consists of documentation work and documentation service. The penumbral region consists of reprograph work, translation, and use of machinery for retrieval translation abstracting, and indexing. The work of the documentalist lies only in the umbral region with a dash of interest as for top management in the work of the penumbral region, which belongs to different kinds of technicians. Education for documentalists should cover depth classification subject heading cataloguing, facet analysis of enquiry of reader, administrative subjects such as selection of documents, acquisition circulation work, and intimate familiarity with documentation periodicals and other reference books, layout and fittings and furniture of documentation centre and the national and international organisation for documentation. Recent fundamental research has opened the way for continuing developmental research in the umbral region. The teaching method should be largely discussional and exploratory backed by tutorial work, practical work and study by students colloquium work, and project work. The training should fire the imagination of the students and stimulate faith in them about the unending challenge of documentation as a career and as a social necessity. The current semantic confusion, caused by the use of the term 'Documentalist' both to denote the wing of the library profession working in the umbral region and also the different kinds of technicians working in the penumbral region should be removed.

A MECHANIZED, MULTI-ACCESS DOCUMENTATION SYSTEM USING [illegible]-DDC AS AN ACCESS POINT. Malcolm Rigby, U. S. Weather Bureau, American Meteorological Society, Box 1736, Washington, D. C.

Limited-use documentation systems may be built around single-access or possibly double-access concepts, but systems designed for many types of uses or users over a long time-span must not be confined to a limited or dead-end concept, but must plan to use all or, or many of the available types of access and display. A versatile mechanized system has been developed using 3 types of "author" access 3 types of "subject" access, chronological, sequential, serial title, language, form of work, and several other m[illegible] of access points to each citation (with or without abstracts). The system may be used for listing, or indexing, or for card or bibliography preparation, either in categories or as a selective retrieval and printout system. The process is completely automatic in that the page size, line length, page numbers, running heads, spacing, sub-headings, supression of or inclusion of any repetitive material, bibliographic citation, indexes or abstracts of any desired degree of abridgement or expansion, (using an IBM 1401-7090 combination,) can be adopted with no additional programming. The special feature of this system which might be generalized is the inclusion of UDC numbers or groups of numbers which may be used to select or arrange by broad, or narrow or correlated concepts or categories. In combination with subject headings, it may be used for highly specialized selection or for display by precise hierarchical subdivision. Economic technical and bibliographic limitations encountered in a limited sample of 1000 documents already stored and indexed by this system can aid in predicting or solving problems which one might encounter in processing a much larger sample. The system can be used equally well with UDC, DDC or a combination of the two, with only slight adaptation.

ON A THEORY OF S[illegible]JECTS Richard [illegible] Raymond [illegible] El[illegible]tr[illegible], 570 Lexington [illegible], New York, N. [illegible]

[illegible] the [illegible] catalogue in index[illegible] distribution [illegible] documents is developed [illegible] documents, requests and processes in [illegible] theoretically [illegible] subject catalogue [illegible] indexing and retrieval it is possible [illegible] probabilities of retrieval [illegible] relevant [illegible] of the [illegible] stated scope [illegible] interest [illegible] are described in terms of the subject [illegible] collection, the average scope of the [illegible], the average scope of user interest, and [illegible] request to document. System [illegible] the number of dimensions and the [illegible] in the subject catalogue and the average [illegible] category. Operation parameters [illegible] categories used in indexing and the [illegible] used in retrieval. Each category is [illegible] catalogue by a number of words in [illegible]. The theory permits computation of [illegible] documents in a given collection [illegible], the probability of retrieval [illegible] expected number of irrelevant documents [illegible] totally satisfactory relations between [illegible] of retrieval and the number of irrelevant [illegible] obtained through good design and [illegible] the system, but the design and operation parameters [illegible] with the demand on the system. [illegible] average are not computed. They will be [illegible]

EURATOM'S NUCLEAR DOCUMENTATION PROJECT. Loll N. Rolling, European Atomic Energy Community, 51 - rue Belliard Brussels 4, Belgium

The paper describes the information needs of Euratom and its member countries and the aims and purposes of the Documentation Project. The fundamental decisions underlying the Project are explained, the development of the system is described and an outline is given of the present activities including the user service, now in its starting period. Particular emphasis is given to thesaurus problems, to the computer processing of index terms, to retrieval strategy and performances, and evaluation of results.

DEVELOPMENT OF A MEDICAL SPECIALTY RECURRING BIBLIOGRAPHY--INDEX OF RHEUMATOLOGY. Mary Jane Kuhl, guest worker, National Institutes of Health, Bldg 10, Rm 3-N-114, Bethesda, Md., U S A

The Index of Rheumatology is a newly-developed recurring bibliography produced by the Medical Literature Analysis and Retrieval System (MEDLARS) of the National Library of Medicine, and published semi-monthly by the American Rheumatism Association. The Index is a product of a year-long literature analysis project (1). "Rheumatology" was defined and delineated according to the results of a survey among rheumatologists. A Thesaurus of Rheumatology was prepared to improve the indexing of publications and to codify the vocabulary. The Index of Rheumatology was developed in cooperation with MEDLARS. This report describes preliminary testing of MEDLARS' output, modification of MEDLARS' Medical Subject Headings (MeSH), and the computer formulation for retrieval of the bibliographic citations comprising the Index. a) Rheumatologic MeSH terms were searched thru MEDLARS. The retrieved citations were examined or correlated for refinement of future retrievals. b) A manual search thru specific issues of journals known to be stored in MEDLARS indicated additional necessary vocabulary modifications. c) Subject headings of peripheral interest were tested by requesting small demand searches from MEDLARS, until the most desirable correlations were discovered. d) New rheumatologic terms were introduced into MeSH. e) The computer formulation was modified and introduced into MEDLARS' recurring bibliography scheme. The Index, which began publication in January 1965, has a present subscription of 750. It is expected to provide over 6000 bibliographic citations annually from MEDLARS' store of periodical literature.

1. Kuhl, M. J., and Sokoloff, L. The literature analysis project of the American Rheumatism Association. Arthr. Rheum. 7: 615, 1964.

(This project was supported by Grant No. AM-07108-01, National Institutes of Health, to the American Rheumatism Association.)

AUTOMATIC INDEXING OF A SCIENTIFIC ABSTRACTS JOURNAL BY THE UDC WITH UNIDEK. Martin Russell, American Geological Institute, 1444 N St., N.W., Washington, D.C., 20005, U.S.A., and Robert R. Freeman, American Meteorological Society, P.O. Box 1736, Washington, D.C., 20013, U.S.A.

The development for Meteorological and Geoastrophysical Titles of UNIDEK (acronym suggesting UDC index), a computer arranged title index based on the Universal Decimal Classification (UDC), suggested the possibility of high-speed, low-cost subject indexes for Geoscience Abstracts. User reaction was sought by issuing both UNIDEK and KWIC (Key Word In Context) indexes with one monthly issue of Geoscience Abstracts. The subject index provided for the full volume consists of a UNIDEK supplemented by a systematic listing of the major index headings and an alphabetic listing of all terms identified by the UDC numbers assigned to abstracts throughout the 1964 volume. Adoption of the system requires UDC schedules, keypunch equipment, a computer capable of handling IBM 1401 programs, and editorial assistance. Advantages of adopting the UNIDEK system include the use of an internationally known indexing vocabulary freely convertible to nearly a dozen major languages and the availability of computer programs already developed. A useful byproduct is the buildup of stored information on tapes or cards, lending itself to mechanical information retrieval searches and compilation of cumulative indexes at longer intervals. Disadvantages (at the time) include the need for updated and fuller UDC schedules, lack of a complete library of UDC schedules on cards or tape, and a general unfamiliarity of the UDC to North American scientists. Full documentation for the UNIDEK computer program is available.

UDC 001.815:681.322.06:025.45

DOCUMENTATION TRAINING IN INDIA. J. Sana, Indian Statistical Institute, Calcutta 35, India.

There is no general agreement on the definition of 'documentation' and the professional image of the documentalist or the information scientist is not well defined. A new figure, however, has come up to integrate scientific and technical information with different stages of activity in scientific research. There is homogeneity in the pattern of information flow through the information systems of individual countries under widely. This is reflected in recent training courses organized in USA, USSR and other countries. Variation within the spectrum of pattern of training according to the needs of the country is logical. In India, the Documentation Research and Training Centre (DRTC) established in 1962 is offering comprehensive instruction on theory and practice of documentation and the curriculum covers universe of knowledge; history of science; depth classification; library catalogue; research and technical library system; documentation; translation service; reprography, and mechanised information processing and retrieval. The one-course apprenticeship is devoted to orientation in library science with particular reference to special library methods and the two-course apprenticeship to reprographic techniques and mechanised information retrieval. The trainees also complete two projects on literature search. Since 1964 the Indian National Scientific Documentation Centre (INSDOC) is also offering one year training on documentation and reprography. While more or less similar, the training at DRTC lays more stress on depth classification and that at INSDOC on reprographic techniques. The training more or less corresponds to that of science librarian and technical literature analyst as defined at the conference on training science information specialists in 1962.

A PROGRAMME FOR BETTER ORGANIZATION OF DOCUMENTATION. **Dr J. Samain, 74, rue des Saints-Pères, Paris, France.**

The Symposium organized by A.I.D. (Association Internationale des Documentalistes) on the "adaptation of the document to its function" has reached the conclusion that cooperation is necessary between documentalists and publishers or producers of documents. New techniques such as reprography, dissemination, retrieval, and regrouping of documents are powerless to satisfy the increasing demand for information if the documents are not designed and presented, from the very beginning, as a function of these requirements and techniques. This cooperation could take place within a large federation grouping documentalists associations and publishers associations, in which the problems raised by standardization, distribution, publicity, copyright and international language would be solved on a practical basis. This cooperation could lead to a flexible and efficient organization, which is described and which is based on the following principles :

- all documents would be published with a complete abstract in English on the first page with reference, author, address, title, complete summary, general line, key words and key ideas.

- each abstract can then be photographically recorded onto microfiches (Filmorex type) so that it can be filed, retrieved and regrouped electronically or visually.

- the abstracts filed or regrouped on microfiches can be disseminated by means of photographic process (Bibliorama, photolisting).

-general or specific indexes could be formed by means of computers (IBM, Bull, ...) from photographically grouped abstracts and from the recorded key-words.

IDENTIFICATION AND CONTROL OF VARIABLES IN INFORMATION RETRIEVAL EXPERIMENTATION. Tefko Saracevic and Alan M. Rees, Center for Documentation and Communication Research, School of Library Science, Western Reserve University, Cleveland, Ohio 44106.

Experimentation in information science has suffered from lack of definition and deficiencies in experimental design and control. With the absence of such control it is often difficult to determine what has in fact been tested and under what conditions. The reproducibility of experiments, if desirable, is consequently open to doubt. Inadequate attention has been given to experimental design with even less care devoted to the provision of procedures for the successful execution of experiments. This paper identifies the essential elements necessary for valid and reliable experiments in the testing of retrieval systems. The variables (components and sub-components) operating within the context of retrieval systems are enumerated, defined and discussed. Models, experimental design and methods of controlling variables in actual experimental situations are illustrated. The meaning of "control" in retrieval experiments is discussed. Attention is given to sources of experimental bias which contaminate research findings. Specific points are discussed with reference to experimentation currently being undertaken in the Comparative Systems Laboratory at Western Reserve, under the sponsorship of the National Institutes of Health (PHS Grant FR-00118-02). Variables and sources of bias are those identified and manipulated in experiments performed on the basis of an IR system model incorporating components related to purpose: discipline, users, file size, and to functions: acquisition, input source, indexing language, coding, file organization, question analysis, search procedures and dissemination. Each of these components is analyzed separately and numerous variables operating within the components are grouped into those associated with system mechanics, human factors or human-system interaction. Since a multiplicity of variables associated with these components operate within any retrieval system it is argued that any experimental design, as a minimum, must take into account the formulation of adequate procedures regarding the control of all the enumerated variables.

DIE GEGENWARTIGE STELLUNG DER DK IN DER WELT. Martin Schuchmann, Deutscher Normenausschuss, Berlin 15, Uhlandstrasse 175.

Die Dewey-Classification[1] kam um die Jahrhundertwende nach Europa und wurde von Otlet und La Fontaine im Institut International de bibliographie (IIB), Brüssel, angewendet. Eine 1. französische Übersetzung mit den neu geschaffenen Anhängezahlen und mit der damit verbundenen neuartigen Methodik des Klassifizierens erschien 1905. Die 2. Brüsseler Ausgabe zwischen 1927/29 ist in den Unterteilungen wesentlich weitergeführt und entsprach dem damaligen Stand der Wissenschaften, sie zeigt unverkennbar die Merkmale einer eigenen schöpferischen Leistung (Otlet, La Fontaine, Bradford, Donker Duyvis, Walther).

Die aus dem IIB hervorgegangene FID widmete sich seit etwa 1930 intensiv der Entwicklung der DK, unterstützt von den nationalen Mitgliedern der Organisation. Aus der idealistischen Arbeit einer kleinen Gruppe entwickelte sich eine gut organisierte internationale Gemeinschaftsarbeit. Anpassung der DK an den Stand der Wissenschaften unter Rücksichtnahme auf die Benutzer ist ständige Aufgabe der FID und im besonderen des Central Classification Committee (CCC). Die DK ist heute das international anerkannte Ordnungssystem der Dokumentation.

Die 2. Brüsseler Ausgabe ist Basis für zahlreiche DK-Gesamtausgaben: Deutsch ab 1934, Englisch ab 1943, Japanisch ab 1951, Spanisch ab 1955, Polnisch ab 1959, Portugiesisch ab 1961, Russisch ab 1963. Daneben bestehen DK-Kurzausgaben in vielen Sprachen und DK-Fachausgaben für zahlreiche Gebiete[2].

[1] Melvil Dewey: Classification and Subject Index for Cataloguing and Arranging the Books and Pamphlets, Amherst, Mass., 1876.

[2] FID Publications Catalogue, FID-Nr.368a, 1965.

SOLUTIONS FOR PROGRAMMING DOCUMENTATION PROBLEMS. Peter Savides, Systems and Procedures Department A05, Systems Development Division, International Business Machines Corporation, Poughkeepsie, New York.

Adequate documentation is a requirement of every released computer program which can be met in different ways according to program use. In a Systems and Procedures department, documentation problems exist because of the need to communicate with several groups each being on a different technical level and requiring diverse documentation formats. Documentation requirements present a challenge to the professional writer; to the programmer it becomes an almost intolerable burden. He must use a set of skills a bit removed from the precise symbol-manipulation and creative, mathematical logic required of his job. Furthermore, the main documentation burden is placed upon him at a bad psychological moment--when the programming job is done and the planning for the next job is underway. The programmer looks upon documentation as a tedious, anticlimactic chore to be dispensed with quickly. Three activities form the basis of the methods we have devised to solve the problem. First, a Documentation Handbook has been written which contains explanations and examples of all the documentation elements required with each released program. Formats and forms are illustrated, and in some cases model paragraphs which need only the change of parameters to be included in a documentation package. Second, the Systems Control group reviews the documentation associated with each program, refusing to assume maintenance responsibility for any program with insufficient documentation. Because most programmers are anxious to be rid of a finished program, meeting this requirement of the Systems Control group provides the motivation to complete the documentation job. Third, a Documentation Specialist assists the programmer in filling his documentation needs.

EVALUATING INDEXING BY REFERENCE TO TERM-CHOICE PATTERNS OF A CRITERION GROUP. Claire K. Schultz and Richard H. Orr, Institute for Advancement of Medical Communication, 4040 Locust Street, Philadelphia, Pennsylvania 19104, U.S.A.

An attempt to arrive at an objective figure of merit for the quality of author indexing led to the development of a method in which sets of terms supplied by authors for a group of documents were scored by how well these "author sets" corresponded with "criterion sets" that were established by merging all the terms chosen independently by each member of a criterion group to characterize the same documents. In the original study, author indexing and indexing derived from document titles ("Comparative Indexing: Terms Supplied by Biomedical Authors and by Document Titles," American Documentation, in press) were assessed relative to the term-choices of a group of research workers that constituted a sample of the potential users of the indexing to be evaluated; however, the general method appeared to be suitable for assessing other types of indexing, such as term assignment by professional indexers or by computer programs, and for using criterion sets established in other ways, e.g., by a group of trained indexers. Because of the method's potential for wider application, the present study explores some of the variables that affect its reproducibility and practicality. The findings to date indicate 1) that assessments of relative merit based on this method are reproducible (a) with different criterion groups drawn from the same scientist population and (b) with different document samples from the same universe; 2) that the size of the criterion group, and of the document corpus, needed for reliable assessment is small enough to make the method practical; and 3) that various methods of weighting and scoring term sets give much the same results. Tests of other variables are in progress. In addition to comparing the quality of indexing achieved by different means, possible applications of the method include 1) measuring the indexing ability or performance of individual indexers, 2) monitoring the output of indexing services, and 3) improvement of authority lists.

PROMOTING RESEARCH AND DEVELOPMENT: UNIVERSITY OF SHEFFIELD POSTGRADUATE LIBRARY SCHOOL'S APPROACH. Herbert Schur, Postgraduate School of Librarianship, University of Sheffield, Sheffield 10, England.

The one-year full-time diploma course at the newly established Postgraduate School of Librarianship of the University of Sheffield equips scientists and technologists with an overall knowledge of the complex intelligence system necessary for promoting research and development. The course is based on the premises that: (1) the science librarian and the scientific information specialist are only parts of a wider system, the scientific intelligence system of an organisation; (2) a more rapid development of such an integrated system promotes research and development, and (3) the system is best staffed by scientists and technologists who thoroughly appreciate all the operations of the system. The selecting of students, content and methods used in the course are outlined.

ASCA--AN AUTOMATIC CURRENT ALERTING SYSTEM BASED ON CITATION INDEXING. Irving H. Sher and Eugene Garfield, Institute for Scientific Information, 325 Chestnut Street, Philadelphia, Pa. 19106, U.S.A.

Automatic Subject Citation Alert (ASCA) is a system for weekly dissemination of data, which selects relevant items from the literature, based on any of several criteria including citation indexing. ASCA bypasses semantic problems involved in formulating questions for language-oriented retrieval systems. Through SPECIFIC QUESTION CITATIONS (any known published work), ASCA will inform a client of the current articles or patents citing any items in his interest profile. The questions in a user's profile are compared with every reference cited in the current literature. In ASCA, the references in the bibliographies of current works serve as unambiguous indexing terms, identifying appropriate current papers. The specificity of citation linkages enables ASCA to select appropriate items from a wide variety of publications including "peripheral" journals which only publish occasional items of interest to a particular user. Other types of questions which may be entered in an ASCA profile include: the REFERENCE AUTHOR (or CITED AUTHOR) QUESTION--which retrieves all current journal items or patents citing any work by a given first author; SOURCE AUTHOR QUESTION--retrieves all current journal items or patents by a given author, regardless of whether he is a primary or secondary author; ORGANIZATION QUESTION--retrieves all current journal items attributed to a given corporation, institution, or other organization; PATENT ASSIGNEE QUESTION--retrieves any current patents assigned to a given firm; PATENT CLASSIFICATION QUESTION--retrieves any current patent by the specified class or subclass number. ASCA covers comprehensively, on a current basis, over 1,070 journals and all U.S. Patents. ASCA allows the user to add new questions at any time, as additional pertinent works are discovered. The weekly ASCA reports also provide a measure of pertinence since the reports show which and how many of the questions in a profile are cited by each current paper to which the client is alerted.

THE "LITERATURE EXPLOSION" AND ITS EFFECT ON DOCUMENTATION THEORY. T.V.Scrivenor, Secretary, Commonwealth Agricultural Bureaux, Farnham House, Farnham Royal, Buckinghamshire, England.

The "literature explosion" is an article of faith with documentalists. It is often asserted that the number of scientific publications is increasing exponentially. Scientific literature is undoubtedly growing, and in certain fields it may even be increasing exponentially; but this is not true of all fields, and in some the rate of growth seems to be relatively slow. Growth can be measured only in relation to the number of publications in circulation, and estimates of this vary from 30,000 to 100,000. It is important to establish and maintain an agreed figure so that the rate of growth in the various fields may be measured accurately. Because it has been assumed that the rate of increase is explosive, elaborate and expensive systems are devised to deal with a situation that in some areas does not exist. These systems emphasize "current awareness" at the expense of depth of coverage. But informative abstracts are still needed by working scientists, especially in developing countries, and the abstract journal is still an effective vehicle of scientific information. It would not be easy to compile a universally acceptable "world list": it would involve agreement on what constitutes a scientific publication and when its death may be presumed, and this would demand a degree of practical international co-operation rarely achieved except in war. It is, however, a task that could be tackled jointly by UNESCO and FID.

RECURRENT BIBLIOGRAPHIES AS CURRENT AWARENESS TOOLS FOR PROBLEM-ORIENTED RESEARCH: EVALUATION OF AN EXPERIMENTAL PRODUCT OF MEDLARS. Andrew M. Sherrington and Richard H. Orr, Institute for Advancement of Medical Communication, 4040 Locust Street, Philadelphia, Pennsylvania 19104, U.S.A.

Advanced documentation systems, such as the Medical Literature Analysis and Retrieval System (MEDLARS) of the National Library of Medicine, can periodically provide bibliographies of current literature tailored to the special interests of relatively small groups of scientists working on a common problem. Seventy scientists engaged in research related to cerebrovascular disease constituted a test group for a prototype service, the Cerebrovascular Bibliography. Issues of this bibliography contained all the citations that had appeared during a 4-month period under 101 selected headings in Index Medicus. The utility of this publication for searching and for current awareness was assessed by questionnaires and use-diaries completed by the test group. This report deals only with data on the potential value of this publication as a current awareness tool. These data include 1) the headings under which a scientist expected to find citations relevant to his work before he became familiar with the publication; 2) in one issue, every citation that, on the basis of the information provided, appeared to be relevant to his own work; and 3) of the citations he identified as relevant, those that were "new" in that he had not previously known of their existence. The data were analyzed 1) to assess the efficiency of the publication in serving the group and each of its members, 2) to develop group and individual relevance patterns, and 3) to characterize reading habits. Among the more interesting findings are the following: The test issue contained 3052 citations of different articles in about 800 different journals; 45% of the articles were in English. The papers relevant to more than 1/3 of the test group could all be found under 16 subject headings, which accounted for 1/3 of the total pages. On an average, each scientist found that about 7% of all the citations were relevant; whereas, under the headings he had, a priori, expected to include relevant citations, about 20% proved to be relevant. Of the relevant citations, some 80% were new to him.

INFORMATION RETRIEVAL AT THE FOOD AND DRUG DIRECTORATE, DEPARTMENT OF NATIONAL HEALTH AND WELFARE, OTTAWA, CANADA. M. Skulski and A.B. Tennenhouse, Food and Drug Directorate, Department of National Health and Welfare, Tunney's Pasture, Ottawa, Canada.

The Canadian Food and Drug Directorate's needs for a central information system and the methods used to fulfill its requirements are described. These consist of a uniterm coordinate indexing system in conjunction with optical coincidence cards for rulings, decisions, advisory opinions, identification of certain solid dosage pharmaceutical preparations, drug plant inspection reports and selected literature references. An edge-notched card system is used for analytical data and food plant inspection reports. Future expansion into a microfilm retrieval system is briefly discussed.

A METHODOLOGY FOR THE ANALYSIS OF INFORMATION SYSTEMS. **David B. Sparks, Mark M. Chodrow, and Gail M. Walsh,** **Information Dynamics Corp., 80 Main St., Reading, Mass. 01867**

The management of scientific and technical information is of fundamental importance as an element of the administration of research and development. Questions of information system design relating to user characteristics, costs, manpower, flow patterns, etc., are, from a management standpoint, at least equal in importance to the technical aspects of information handling and its detailed organization for any specific purpose. In the past, these problems of systems design have largely been left unexplored by those active in documentation and information sciences, or have received only general empirical treatment. Modern scientific research, however, is recognized as far too important an endeavor, and its information needs far too complex to allow the development of large-scale supporting information systems to proceed without adequate and scientific planning and testing of alternatives.

The work reported in this paper is an initial effort to assemble the elements of information system design as a problem in scientific management, and to propose a methodology for its solution. The application of this methodology involves the analysis of the features of science information systems and the identification of the interrelationships among them. Techniques for a structured mathematical representation and quantification of these features and relationships were developed and reduced to computer practice. Simulated data to test the methodology were developed from current statistical sources and a mathematical model of a range of alternative information system configurations were constructed. The mathematical calculations in the models were performed by computer, and the results analyzed, for the purpose of effecting meaningful comparisons of system alternatives.

This work was sponsored by the National Science Foundation under contracts NSF-C-263 and NSF-C-370.

MATHEMATICAL ANALYSIS OF DOCUMENTATION SYSTEMS' Dagobert Soergel, 78 Freiburg, Rotebuck 68, Western Germany

As an attempt to a general structural theory of information retrieval a documentation system (DS) is defined as a formal system consisting of a) a set (a) of objects (documents); b) a set (A^+) of elementary attributes (key-words), from which further attributes may be constructed: (A^*) generates (A); c) a set of axioms of the form $X(x) = m$ ($m \in M$, M a set of constants) connecting attributes with objects; from the axioms further theorems (= true sentences) may be constructed. Using the theorems, different mappings (A)→(a) (search question→ set of documents retrieved) are defined. The type of a DS depends on two basic decisions: 1) choice of M: M may consist of the logical constants "true" and "false", or of some positive integers etc.; 2) choice of the rules for the construction of attributes and theorems; e.g. logical product in coordinate indexing (CI); links. Further practical decisions: (A^+) hierarchical or not; kind of mapping, introduction of roles (= further attributes). The most simple case - ordinary two-valued CI - is discussed in detail: (A) is a free distributive (but not Boolean) lattice, the homomorphic image a ring of subsets of (a); a useful retrieval operation "praeternegation" is introduced. Furthermore are discussed: a generalized definition of superimposed coding; some functions for the distance of objects or attributes; optimization and automatic creation of classifications. The model may be extended to take into account term-term- and document-document-relations. It may serve as a structural frame in terms of which the functional problems of retrieval theory may be expressed more clearly.

Acknowledgements to Prof. Dr. J. Nitsche, Institute of Applied Mathematics, University of Freiburg.

COORDINATION OF INFORMATION STORAGE AND RETRIEVAL SYSTEMS BY MEANS OF THE TS 15 SYSTEM. Jiří Spirit and Vladimír Zhannabauer, Institute for Technical and Economic Information, Konviktska 5, Praha 1, Czechoslovakia.

The development of information storage and retrieval system leads to a number of projects in the various fields of natural, technical and social sciences; the main source of differentiation is the variety of the intellectual treating (=analysis and indexing) of the document and "the software" in each scientific discipline, while the "hardware" could be unified in all branches. Cooperation between information centers in the different domains suffers at the time from lack of adequate means of communication, especially of mediums for information storage which would be interchangeable between the individual information centres. The TS 15 information storage and retrieval system divides human knowledge in several hundred branches which are specified by a code number; the individual documents, too, carry a code number of the appropriate branch and the document's sequence number in that branch. Documents are analyzed and indexed according to the types of data and roles specified for that branch. These data are punched into the card, sorted and tabulated. The information storage process results in a group of tables with various data sequences. The sequence number obtained from the combination of data (corresponding to the request) equal the sequence numbers of the abstracts and original documents. The TS 15 system thus enables various indexing systems to be used for various branches and to use unified techniques and technological processes for all scientific disciplines.

FACTORS INFLUENCING INFORMATION DISSEMINATION IN SCIENCE AND INDUSTRY, Jiří Spirit, and Vladimír Thannabauer, Institute for Technical and Economic Information, Konviktska 5, Praha 1, Czechoslovakia.

When choosing the right organization, methods and means for information dissemination in science and industry one has to decide first of what this importance this information is to the manager, scientist, designer or engineer. The value of this information for any possible user should be determined by a formula containing all factors concerning the user's profession, branch of information activities, type of document, form of dissemination and the point of view of information. This formula serves both for disseminating current information and individual retrieval. Each factor is defined by a combination of circumstances mentioned above and its value equals the root of 0,1 2 or 3 raised to the n-th power. The value of the root depends on a given combination of circumstances, the value of the exponent depending on the relative importance of the chosen circumstance as such. The value of the factors are printed in tables. The report contains a survey of various forms of disseminating scientific information and practical examples of using such information in management and creative processes.

DEVELOPMENT OF A PROGRAMMATICAL - TECHNICAL SUBJ[E]CT INDEX FOR THE OFFICE OF AEROSPACE RESEARCH. Robert W. Summers, Chief, Physics Branch and Dr. Monroe Freer Director, Science Information Exchange - Smithsonian Institution, 1730 M Street, N. W., Washington, D. C., U. S. A.

The Exchange has developed a novel subject index for the Office of Aerospace Research "Air Force Research Résumés" Volumes 4 and 5. Methods [illegible] comprehensi[ve] subject index were explored and tested. The problems encoun[tered] were the index format, [illegible] to be associated with a given index point, those associated with analys[is] of technical content, and the [illegible] of selected infor[mation] [illegible] processing. Index format and the selection of information established the basis for relating mechanics of technical analysis and [illegible] handling. Each ind[ex] point is a composite of one to three categories and their corresponding interrelationships indicated. Paraphrased sentences were synthesized to expand the composite index point to give maximum intelligence to that point. The development of the index progress[ed] stepwise through reviews, edits and analysis of the index. While the index is rather unique in appearance, it seems to be fairly com[prehensive] in function. Permutations of the "key words or phrases" are somewhat unconventional creating modifiers, qualifiers and generic descendents on a hierarchical scheme. Neither the terms nor their permutations were preassigned, but were developed progressively from the subject content by staff scientists as appropri[ate,] consistent and useful. The index affords a complementary combi[na]tion of fields for program management, as well as the selection of specific topics in technical detail.

[...] DEL MEDIO Y APLICACIÓN DE LA CLASIFICACIÓN DECIMAL UNIVERSAL EN LA REPÚBLICA ARGENTINA. Reinaldo José Suárez, Biblioteca Central de Marina, Bartolomé Mitre 1405, Buenos Aires, Argentina.

Compilación de datos e inclusión de informes, estudios especiales, colaboran en los mismos: Raúl Cisneros Malbrán (Córdoba); Demetrio Sagastizabal (Rosario de Sta. Fé); María Julia Fernández de Loureiro (Buenos Aires); Esther M. Ferraté (Entre Ríos); Paulina D. Giovanaz (Buenos Aires); Enrique Kreibohm (Tucumán); Ricardo J. Lois (Buenos Aires); Nicolás Matijevic (Bahía Blanca); María Colteni (Buenos Aires); Daniel Roque (Córdoba). Se analiza su evolución a través de los trabajos, a principios de siglo, del Dr. Fernando Lahille (1861-1940), continuados por el Ing. Federico Birabén (1867-1929) y por el Prof. Pedro B. Franco (1894-1947).- Se comentan los aportes de Manuel Selva y la actuación de José Hartmann (Hospital Ferrari-Academia Nacional de Medicina). Se detiene en la acción cumplida por diversas instituciones y bibliotecas en pro del sistema: Bibliotecas de Marina, Centro de Estudios Bibliotecológicos (Comisión de estudios de la CDU), Dirección Nacional de Vialidad, Centro Argentino de Ingenieros, Biblioteca Popular del Municipio, Comité para la CDU (Consejo Nacional de Investigaciones Científicas y Técnicas), Instituto Bibliotecológico (Universidad de Buenos Aires), ex-Oficina Bibliográfica y Biblioteca Mayor (Universidad Nacional de Córdoba), Biblioteca Nacional de Aeronáutica, Biblioteca Central (Secretaría de Estado de Agricultura y Ganadería), Centro de Documentación Bibliotecológica (Universidad Nacional del Sur, Bahía Blanca (B.A.), Biblioteca Central (Facultad de Filosofía y Letras, Universidad de Buenos Aires), etc. Incluye referencias bibliográficas y listas de los elementos con que cuentan algunas bibliotecas para las tareas de clasificación.-

FLEXIBILITY PLUS - THE IBM 870 DOCUMENT WRITING SYSTEM

Karen G. Takle and Burton E. Lamkin

Systems Development Division Library
IBM Corporation
San Jose, California

Some of the significant methods by which the IBM 870 Document Writing System can speed acquisition, processing, distribution, reference and educational services in a librar[y] are illustrated in this paper. The IBM 870 features flexibility and economy particularly as it relates to library applications. Each 870 system consists of operating units which can be divided into three functional groupings (1) inp[ut] devices,(2) a control unit;(3) output devices. Data entered an internal transmission circuit called a "common channel" is available to the output units for the concurrent preparat[ion] of one or two punched cards, two typed documents and eith[er] a 5-track or 8-track paper-tape record. No two of these output documents are required to use the same data or document format. Because of the great flexibility of this machine, complicated data formats can be accommodated [by] this system. Current library applications of the IBM 870 Document Writing System are discussed and examples of flexibility and applicability are drawn from current operat[ing] systems. Flow charts illustrate the key role of the 870 system in the IBM Development Laboratory Library at San Jose, California. Discussed are multiple input and output formats which illustrate the practicability of this system.

A MECHANIZED REGISTRY OF CHEMICAL COMPOUNDS F. A Tate, H L Morgan, D P Leiter, & R E Stobaugh, Chemical Abstracts Service, Columbus, Ohio

A computer based system has been established for storing and retrieving the structures of chemical compounds along with citations to literature describing the compounds. The system, called the Registry System, includes a set of interrelated files which record a machine language description of the structural diagram, the molecular formula, accepted nomenclature, and bibliographic references. Registration (entry into the System) includes assigning a unique machine address (Registry Number) to each compound new to the System or retrieving the previously assigned Registry Number from the files. It is intended that the System ultimately include a record of every chemical substance reported in the literature and identification of all useful published literature bearing on each substance.

Computer records are being prepared for all compounds which are reported and adequately identified in the primary journals from which abstracts will have appeared in Chemical Abstracts since January 1, 1965. Special CAS reference files will also provide compounds for input to the Registry System. Four means of computer file build-up are planned through the registration of all compounds for which structures are drawn in indexing Chemical Abstracts, frequently reported compounds each of which may be reported under a variety of designations, compounds included in previously published CA indexes through a name matching process, and special collections of compounds selected from CAS indexes. This four-way approach has been designed to provide the most economical and rapid route to a complete operating system.

Initial sources of input have been chosen to achieve both maximum growth rate of computer files and the widest range of operational and technical challenges for the System. A major objective of initial work is to accumulate economic data under operating conditions. Another early objective is the development of literature analysis tools through the use of which staffing requirements for the System may be maintained within practical limits. This work is partially supported by funds from the Department of Defense, the National Institutes of Health, and the National Science Foundation through a contract with the National Science Foundation.

IMPRECISION IN MEANING MEASURED BY INCONSISTENCY OF INDEXING. Dr. John F. Tinker, Eastman Kodak Company, 343 State Street, Rochester, New York, U.S.A.

Meaning is defined as the relevance of a word to the concept it labels. Following Hillman (Proc. of Amer. Doc. Inst. meeting Oct. 64, v.1, p.3) meaning is tetradic: the relation of a word and a concept, to a degree, within a corpus. Indexers, in choosing or assigning all words strongly associated with concepts of a document, assert that the document means the word; therefore, consistency of indexing measures the precision with which meaning is understood by the indexers. From an indexing experiment designed to avoid all judgments of truth or correctness of indexing, the following results were obtained: 1) When descriptors are freely chosen, the more descriptors assigned to a document the more difficult is the retrieval of it. 2) By restricting the descriptors used in searching, the disadvantage noted in (1) can be overcome. 3) Possibly, older words are used more frequently and less precisely than newer words. 4) A graph of the total number of applications of a descriptor vs the ranked document serial number lucidly illustrates the precision of meaning of the descriptor; a precisely understood word gives a rectangular curve, a less precise one an S-shape. Many words have imprecise meanings, even to specialists in scientific fields.

COMPARISON OF CLASSIFICATION SYSTEMS IN THE FIELD OF CIVIL ENGINEERING. João Fernando Cansado Tavares, Núcleo de Documentação Técnica, Rua Gonçalves Zarco, Restelo, Lisboa, Portugal.

After stressing that the problem of classification is basic in every information department, the author briefly mentions the most used traditional systems and the modern information retrieval methods. The classical U.D.C. systems and the automatic methods based on "Keywords" operated by optical means or punched cards are described. Applying both systems to an instance, the author compares their efficiency, concluding that the advantages of U.D.C. vanish in fields requiring great detail. Finally some conclusions are drawn which can influence the choice of classification system, in function of the following parameters: field of knowledge covered; number of references to be classified; extend of detail required; type and frequency of information demands.

PHYSICS THESAURUS PROJECT. Allin Cyrulewicz, Physics Abstracts, Institution of Electrical Engineers, London, U.K. and Pauline Atherton, Documentation Research Project, American Institute of Physics, Washington, D.C., U.S.A.

The object of the project is to produce a physics thesaurus starting from the present alphabetical subject index of "Physics Abstracts", in order to (i) enhance the usefulness of the subject index of "Physics Abstracts" by improving the ease of access to the indexes, the users and increasing the uniformity of indexing by the editors; (ii) provide the basis of a system for alphabetic, coordinate and other methods of indexing physics material by organizations other than "Physics Abstracts", (iii) provide the starting point for a more extensive interdisciplinary thesaurus on a national or international level. The first stage of the project has now been completed by the editorial staff of "Physics Abstracts" at the Institution of Electrical Engineers in cooperation with the Documentation Research Project of the American Institute of Physics. This stage involved the following changes in the subject index of "Physics Abstracts": (i) adding several hundred "see also" and "see" references; (ii) defining, where necessary, headings and subheadings by scope notes; (iii) collecting all the "see also" and "see" references from other indexing points under the headings to which they refer; (iv) standardizing the first word (s) to be used in entries under the various headings; (v) coding each term according to the chapter arrangement for abstracts in the monthly issues of "Physics Abstracts", thereby providing for automatic grouping of terms, cross references, scope notes, etc. for given fields of physics. The new version of the subject index resulting from these changes was analysed for relationships between concepts indexed and further alterations were made in the light of these relationships. The draft physics thesaurus produced by these operations is available in machine readable form. Copies will be submitted to a selected group of physicists for critical comment. The comments obtained will be acted upon, to establish a new physics thesaurus and alter accordingly the subject index of "Physics Abstracts".

THE NECESSITY OF COMBINING SELECTIVE DISSEMINATION OF INFORMATION AND INFORMATION RETRIEVAL WITH MACHINE TRANSLATION. P.E.Vásárhelyi, Scientific Institute for Technical Information, Budapest, Hungary.

In Hungary Russian and English language abstracts are often the base of information work. The Scientific Institute for Technical Information is working now on a project in order to develop a method combining the mechanization of several steps of information work with machine translation. The proposed system works as follows:
The important words in the Russian/English abstract are punched into tape and read in the computer. A Russian-Hungarian and an English-Hungarian dictionary of a special subject field are stored on magnetic disks and the significant words of the abstracts are matched against the dictionary. The computer assignes the Hungarian equivalent to the word /eliminating at the same time the synonyms/ and it is now the Hungarian word which is compared with the profile of the users in order to find the persons waiting for information on a special subject. The user gets the bibliographic reference with all the significant words of the abstract translated into Hungarian. The data remain stored in the computer forming thus a document file that may be searched for documents containing specific information.

SYSTEMS CONCEPTS FOR ESTABLISHING AN AUTOMATED INFORMATION CENTER. Vincent J. Vitagliano, IBM Corporation, 555 Madison Avenue, New York, New York, U.S.A.

The introduction of automated information services at Engineering Index is used as a case history for discussing the systems concepts for establishing automated information centers. The installation represents a joint effort by Engineering Index and the American Society for Metals in conjunction with Battelle Memorial Institute and the IBM Corporation. Recognition of professional areas, attitudes and objectives are considered to be prime factors for a successful transition from manual to machine operations. The conversion of intellectual efforts to machine readable forms, and the aspects of programming, programs, and computer operations are presented. A generalized combined file search computer programming system for the IBM 1401 Data Processing System has been used to create a dictionary, a master file, and an inverted file, and to perform retrospective searches. In addition a series of publishing programs utilizing the master file produces subject heading and author indexes. The dictionary is used for descriptor control, whereby input to the system is checked for validity and correctness. The master file contains the information that is pertinent to the documents which have been abstracted and indexed by a professional staff. The inverted file lists each descriptor followed by the numbers of the documents to which that descriptor has been assigned.

THE ROLE OF THE BOOKSELLER. W.I. Veasey, B.H. Blackwell Ltd., Broad Street, Oxford, England.

The bookseller has an important role to play in the organisation of information. The information must be at hand, borrowing is frequently difficult or undesirable; and for the procurement of his material the librarian must rely upon the professional skill, experience and perseverance of his bookseller. Anybody can buy the latest item on a well-known publisher's list, but much more is required if the documentalist is to obtain for his use the obscure pamphlet, the haphazard periodical, the unadvertised report. Booksellers' standards vary from country to country and firm to firm, but documentalists everywhere require the highest standards. These can only be attained in certain basic conditions are fulfilled. There must be fundamental willingness on the part of the bookseller to attempt to obtain any published item for which he received an order, and to this end he must know and record the habits, customs and idiosyncracies of publishers of many lands. He must be prepared to train, and pay, staff to be professional booksellers, not just counter assistants; and the training of a good bookseller is as stringent and demanding as that of any librarian. He must be prepared to disseminate information on publications, new and old, by traditional methods such as catalogues down the latest machine-recorded profile of subject interests of his customers. Booksellers should also be efficient business men: they need to be if they are to stay in business. And for his part the customer must accept that his bookseller is an expert working towards the same ends as himself, and not expect a first class service at a second class price. There is an increasing need for greater co-operation and understanding between librarians and their booksellers.

[illegible] DIFFERENT TYPES OF CLASSIFICATION SYSTEMS [illegible] THEIR OBJECTS. Ingar Ahlin, Sturegatan 102, Stockholm, Sweden.

This paper deals with fundamental problems in classification research, some questions to which prominence was given in the recommendations of the FID/CR Conference in Elsinore 1964 are discussed. The main line is to analyse different categories of classification systems characterized by different objects.

The systematics of science as a necessary but neglected general background for studies on other systems.

Document systems are developed from systems of sciences (the most important systems from an now antiquated model). As documents represent combinations of concepts, doc. systems have to indicate positions for the most important concept combinations. The hierarchical structure therefore must be partly non-generic. Same terms are found in different places; no completeness in respect of terms is aimed at. Possibilities and limitations of doc. systems are discussed. The principles for UDC as given by UDC seems to be contradictory and the recommended combinations between already complex subjects leads to an incomprehensible situation. Guiding lines for developing a new universal system are advanced.

Term systems. Coordinative indexing actualizes the systematical arrangement of terms. Some thesauri from different branches are investigated. The compatibility in overlapping fields is studied. An systematical approach to the thesaurus problem is necessary, but the hierachical structure in term systems can't be the same as in doc. systems. Generic thesaurus works in special fields can't be regarded as contributions to a universal thesaurus, as the terms in these special thesauri are looked at from a special point of view.

Systems for technical products are given a lot of attention in all industrial branches. The relation between systems of this kind and other systems is discussed.

POSSIBILITIES OF ARTICULATION OF INFORMATION CENTERS INTO A NETWORK. Eugene Wall, Information Dynamics Corporation, 2000 K Street, NW, Washington, D.C., U.S.A.

Although there are only eight basic types of information systems (based upon the possible combinations of internal functions which systems can perform), it is shown in this paper that the interconnection of systems into a network permits sharing of effort in certain respects so that 14 types of systems are possible. This is at least true with respect to systems handling serial literature, but it is believed that those handling separate literature will be identical, nearly identical, or closely equivalent. It is shown that each of the 14 system types produces a different combination of outputs for the use of its clientele or for the use of other systems. When the variations in combinations of outputs transmitted between pairs of system types are examined, it is concluded that only 13 different modes of such communication - i.e., "modes of articulation" - are possible.

It then becomes apparent that the 14 types of systems constitute only six overall groups; all system types within a group interact with each other system type in identical manners. Finally, it is shown that the 13 modes of articulation can be organized into hierarchies, with the more general modes subsuming the more specific modes - pointing to the possibilities of (1) planned routes for evolutionary development of systems' articulative capabilities and (2) the need for only a very few standard procedures of articulation from which the more specific modes may be excerpted.

In short, articulation among information systems for handling serial literature is not technically complex on the "system of systems" level (i.e., determining what is to be done) once the requirements of the clientele of the individual component systems have been determined. The matter of determining how to implement the what may well be technically and sociopolitically difficult, and perhaps even economically impracticable, although the theory and practice of all aspects of such implementation have already been individually proved.

THE COMBINED FILE SEARCH SYSTEM: A CASE STUDY OF SYSTEM DESIGN FOR INFORMATION RETRIEVAL. L. A. Warheit, International Business Machines Corporation, Monterey and Cottle Roads, San Jose, California, U. S. A.

A requirement was established to design a general purpose, computer-based information storage and retrieval system. Many types of information files were collected. Their characteristics and the types of use to which they were put were tabulated. These various operating systems were grouped into three broad types: 1) formatted files, 2) indexed document retrieval and information handling systems and 3) unindexed full text processing systems. The most immediate need was to provide a computer program for the second type. In order for it to be really general purpose, the program had to be able to accept files of all sizes, provide for practically unlimited growth, be able to give real time response, be efficient and economical, modular, flexible, and provide for document and data retrieval, i.e. have all the normal Boolean and arithmetic operators and yet be used with a small computer. A program was written to meet these requirements and the system installed for a number of applications. The program will be described in sufficient detail to show how the design objectives were met. Operating experience, which is being continually evaluated, is proving that the Combined File Search system meets all the basic requirements. Based on this experience, the system is being further refined and is being programmed for the new generation of computers.

A STUDY OF SUBJECT ANALYSIS. Jiří Wanke, Institute for Technical and Economic Information, Konviktska 5, Praha 1, Czechoslovakia.

A methodology of subject analysis has been elaborated for use in the Czechoslovak information services. General criteria have been suggested for acceptable abstracts, man-made or machine-made, and for other products of document analysis (selection of descriptors, subject indices). Abstracts are to be distinguished as to their purpose (a) for current awareness (CA) and, (b) for storage and retrieval (SR), rather than to their form (informative and indicative). While for CA abstracts the generally accepted informative form is recommended, special recomendations are set up for SR abstracts, which differ from CA abstracts in the basic approach, style, selection of data, type of documents analyzed, life, date of first use, classification requirements, processing, application aptitude to formalization. In a conventional generic file, SR abstracts may be considered as classification extended ad hoc into areas where the generic structure could not be anticipated and/or made in detail. They are to be regarded as a non-linear transformation of the original, followed by a description of the document itself. The degree of non-linearity should be given for any point so as to enable the user to boost the abstract to the original size of the document. The SR abstract should be made with regard to the contents of all preceding documents in the file and even of documents expected to come. Any incoming information should be confronted with information stored already; all redundant data should be rejected in order to save storage medium. (Machine-made abstracts fail in these points.) Certain types of information (e.g. statistical data, parameters, factografic matter) should be extracted along special rules and stored separately for direct access. In formalizing abstract syntax and standardazing terminology we approach from another side the goal followed by adding to descriptors roles, links, flexions and similar devices; i.e. a general purpose information language.

INDEX PUBLISHING SYSTEM WITH GENERATED CROSS REFERENCES AND SELECTIVE BODY CONTENT. Michael L. Wascak, IBM Corporation, 2925 Euclid Avenue, Cleveland, Ohio U. S. A.

This system was functionally designed to produce flexible indexes for information files of all sizes. As a general purpose system, it is currently producing a Subject Index for each of three different disciplines. Subject headings may consist of single access terms or of access terms with a subordinate term appended. They appear in alphabetical order, followed by document entries. Each document entry consists of a document number and the segments selected at the beginning of each publishing run. Cross references are generated by the computer, from the thesaurus containing all term relationships used in the total system. Book page format with a running head is produced by the computer. Each page consists of two columns of content ready for photo reduction to the size of the publication page. Option of heavy printin, (double impression) each index entry is a system feature. Another index can be produced simultaneously or separately as required. Currently, it is a author index being produced in the same book page format. Input is a serial file of documents, such as in the General Purpose Information Storage and Retrieval System now in the IBM 1401 Library, number 10 3 047.

NASA SDI PROGRAM. Van A. Wente and Gifford A. Young, Scientific and Technical Information Division, National Aeronautics and Space Administration, Washington, D. C., U.S.A.

Directed announcement and document distribution through an automated selective dissemination system are serving otherwise unfulfilled information needs of NASA space scientists and engineers. NASA's developmental Selective Dissemination of Information program (NASA/SDI) was initially developed under contract by IBM Advanced Systems Development Division and is now operated by NASA's Scientific and Technical Information Facility. Over 2000 documents, the full contents of the combined issues of Scientific and Technical Aerospace Reports and International Aerospace Abstracts, are compared twice monthly with the expressed interests of nearly 700 participants. Centralized matching provides announcement service to these users located at 10 NASA and 11 Air Force Centers, with decentralized operation planned for Centers having requisite machine capability. System output consists of tab-card sized abstract cards duplicating photo-composed journal announcements plus a response card, which may be used by participants to request the full-text of selected documents. Provision of requested full-text as microfiche, facsimile, or original document is the responsibility of each Center, with backup assistance by the central NASA Facility. Exceptional flexibility is possible in expressing participants' interests, match options include "must" single-word descriptors, two- to seven-word phrases, "not" terms and phrases, and percentage matching. In preparing the input document and user interest profiles for matching, a dictionary program provides computer-generated codes and can substitute cross-referenced descriptors. The computer programs employ an IBM 7090/94 with 32K storage, two 7607 data channels, and eight 729 tape drives. Other NASA/SDI concepts under test include citation printout rather than reproduced abstracts, and matching by summations of topical profiles rather than by individual participant profiles.

PROJECT CADRE, CURRENT AWARENESS AND DOCUMENT RETRIEVAL FOR ENGINEERS; A STUDY IN PROGRESS AT ENGINEERING INDEX. Fred R. Whaley, Battelle Memorial Institute, and Carolyn Flanagan, Engineering Index, 345 E. 47th Street, New York, N.Y., U.S.A.

Engineering Index has embarked on a program aiming to provide broader, deeper and faster abstracting and indexing services in all the engineering disciplines. As an initial step toward this goal, a pilot study is under way in the electrical-electronics field and the plastics field. A monthly abstract bulletin has been launched in each of these fields as segments of Engineering Index. Eventually a collated index of all the segments will be available. Cumulative indexes will also be compiled. In addition, a deep index employing links and EJC roles is being constructed, which will be the basis of a central mechanized document retrieval service. The present current awareness services of Engineering Index will be extended and refined so as to give selective dissemination of information to subscribers. These combined services constitute an important step toward the long-range goal of establishing an Engineering Information Center at the United Engineering Center in New York. The system is called CAIRE, for Current Awareness and Document Retrieval for Engineers. Journals are scanned and pertinent articles analyzed by professional personnel in accordance with the EJC system of abstracting and indexing. Their work sheets are processed using a computer for both the index print-out for the bulletin and for the machine search and SDI capability. The heart of the system is a thesaurus, which serves not only as a visual aid to indexers and inquirers, but also directs the following computer functions: (1) validate all index terms and subject headings; (2) substitute preferred terms and (3) generate in the published bulletin "see" or "see also" references for terms previously designated for this purpose in the thesaurus. The basic computer programs were developed by IBM, but many special features have been added for this project. The EI thesaurus is based on the Thesaurus of Engineering Terms (EJC), but contains about 30% new terms. The American Society for Metals has adopted the same overall system, and a cooperative program between EI and ASM is in effect.

INFORMATION STORAGE AND RETRIEVAL BY PRE-COORDINATION. C. J. Wessel*, Food and Drug Administration, Washington, D.C., and John Dere, Consultant, Washington, D. C., U.S.A.

A method by which all information stored in a system is permuted in anticipation of possible or organized associations of descriptors provides a printout ready for reproduction, with the answers to possible inquiries. As a result, one has at all times a complete inventory of the stored information, accessible from the various points of view from which this information may be considered.

Aside from the advantage of permitting an inventory look-up for the subsequent search, the printout opens an excellent browsing field guiding the inquirer to synonyms and associations or other useful terms which were not anticipated, without look-up in a thesaurus.

The printout is modular in regard to the various items of the bibliographic citation which can be included. The display of the permutation of descriptors and document numbers may vary according to the type of collection or the printout method selected. This printout may include a display of associated descriptors with all the document numbers common to those descriptors; the occurrence of pairs or multiple associations can be traced.

The specific descriptors assigned to each document or the abstract of this document may be included in the printout.

The bibliographical information may be included in the printout under each descriptor, or may be printed out in a serial manner.

Conclusion: The method consists in a synoptic grouping of indexing and bibliographic information available on demand. Further refinement in association patterns is possible; the technique used is flexible and can be adapted to the changing needs or the size of a document collection.

*This work was performed while this author was associated with the Prevention of Deterioration Center, National Academy of Sciences-National Research Council.

INTERNATIONAL BIBLIOGRAPHY OF THE SACRED SCIENCES. Eugene P. Willging, Catholic University of America Library, Washington, D. C. 20017

Of more than 500 scholarly titles of serials pertaining to theology, church history, canon law, Scripture, and related disciplines, more than 90% are indexed fully or partially in duplicate (up to ten times) in 20 different bibliographical tools. To show extent of coverage and amount of duplication a comprehensive list is scheduled for Sept. 1965 publication. As a by-product a plan is being developed to coordinate the indexing of books and serials in the sacred sciences in a new tool tentatively titled INTERNATIONAL BIBLIOGRAPHY OF THE SACRED SCIENCES, which would appear monthly under interdenominational auspices. Through inclusion of subject headings, classification numbers, multiple entries for editors, translators, etc. with the main author entry for books, it would attempt to solve a major library problem of cataloging backlogs as well as provide theological scholars with quicker indexing as well as more intensive coverage and more frequent cumulations.

TENDENCIES IN THE ORGANISATION OF SCIENTIFIC AND TECHNICAL INFORMATION SYSTEMS IN CERTAIN EUROPEAN COUNTRIES. Adam Wysocki, Documentation and Scientific Information Centre of the Polish Academy of Sciences, Nowy Swiat 72, Warsaw, Poland.

On the background of the actual state of organisation systems and basic forms of activity in the fields of the scientific and technical information an analysis is done of the forthcoming tendencies in the organisation of gathering, working, preserving, searching and disseminating information in Poland and other East European countries. A study of the organisation of science and national economy shows its influence on the ideas of organisation of the scientific and technical information systems in these states as well as the development possibilities of information systems subject to this analysis.

ORGANIZATION OF "CURRENT MEDICAL TERMINOLOGY" FOR COMPUTER GENERATION OF PROGRAMMED TEACHING TEXTS. Hans H. Zinsser, M.D. Susan Graham, and David M. Smith, Columbia University, College of Physicians and Surgeons, 630 West 168th Street, New York, New York, U.S.A.

The availability of streamlined and definitive medical information on magnetic tape derived by the American Medical Association staff has prompted us to manipulate this tape in such fashion as to generate teaching sequences by direct similarity analysis of the work and phrase content of the information. Generation of branched trees of differential diagnostic sequences is followed by the insertion of general statement and question and answer sequences at predetermined iteration rates and with logical and work similarity cues. By this sequence, any body of work-related information can be processed to yield at least a rough draft of programmed text material, in any language. The saving in man hours using the IBM 7094 is of three orders of magnitude.

THE U.S. FOREST SERVICE STUDY OF THE OXFORD SYSTEM OF DECIMAL CLASSIFICATION FOR FORESTRY. Theodor B. Yerke, Librarian, U.S. Forest Service, Berkeley, California, U.S.A.

The United States has no established forestry classification. The Oxford system expands UDC 634.0 (Forestry). In 1961 the Forest Service began to evaluate Oxford's potential for literature control in research and operating units. A catalog of 70,000 Oxford-classed citation cards was built. This tool was scrutinized card-by-card to determine the validity of the system's claim to [illegible] special degree of precision, flexibility, and capacity for growth.

Evaluative criteria included capacity to handle micro- as well as macro-information, to reflect changes in the structure of research, and to avoid Procrustean operations. The validity of related UDC schedules, which must be used when Oxford's scope is exceeded, was also investigated.

The three-year study shows that Oxford handles general and traditional forestry literature adequately. Difficulty is encountered with the recognition of new modes of thought. Oxford is not well fashioned for handling specifics of micro-information. Scheduled revisions to expand for this results in excessive notation. For advanced and pioneer research milieux, it is better to turn to more recent information handling techniques.

An 'Americanized' index has been prepared for the Oxford system. This includes "americanization" of terms and scope notes. The index is planned as a Forest Service Handbook, for general use in Forest Service units.

INFLUENCE OF SELECTION AND WIDE DISTRIBUTION ON TRANSLATION USE. Raymund L. Zwemer, FASEB Translation Project, 9650 Wisconsin Avenue, Bethesda, Maryland, U.S.A.

The Federation of American Societies for Experimental Biology under a contract with the National Library of Medicine undertook this experimental approach to the problem of translation as an alternative to a previous program of complete translation of a limited number of journals. Ten percent of 20,000 research articles in 45 journals of USSR and East European origin were carefully selected and translated into good scientific English. Translated articles were screened by referee editors for novelty of material, statistical validity, concepts, conclusions and citation of current literature. About one third of the translated articles were published in Federation Proceedings Translation Supplement which provided wide distribution to 12,500 subscribers of which 2,800 were libraries. The tenth bimonthly issue for July 1964 carried a simple questionnaire asking whether readers had found one or more articles of interest. The geographic distribution of responses corresponded closely to that of FASEB membership. Distribution by discipline showed that members of one of the six FASEB member societies found less of interest (60%) than did members of the other five (90%). A possible explanation may be the criteria for journal selection. Serials could not be scanned if they were being translated by others. Important biochemical and biophysical journals were thus excluded. Other criteria for journal and article selection will be given and a more detailed analysis of the replies to the questionnaire, including recording and citation of articles. One advantage of our plan of distribution is that complete translations with tables and illustrations are made directly available to nearly 10,000 individual users together with a listing of additional translations available at a small cost on demand. Correspondence shows that the translations are also being made known by a variety of specialized abstracting and indexing services.

INDEX